Fri Jan 19 06:20:01 2001

us-09-294-298-5.rge

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 07:18:54 ; Search time 14302.9 Seconds
                                              (without alignments)
                                              1424.454 Million cell updates/sec

Title: US-09-294-298-5
Perfect score: 3981
Sequence: 1 taggagagactgagctgcc........tgggtgtgtatcccatcct 3981

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
    1: gb_bal:*
    2: gb_ba2:*
    3: gb_om:*
    4: gb_ov:*
    5: gb_ph:*
    6: gb_pl1:*
    7: gb_pl2:*
    8: gb_pr1:*
    9: gb_pr2:*
   10: gb_pr3:*
   11: em_fun:*
   12: em_hum1:*
   13: em_hum2:*
   14: em_in:*
   15: em_om:*
   16: em_or:*
   17: em_ov:*
   18: em_pat:*
   19: em_ph:*
   20: em_pl:*
   21: em_ro:*
   22: em_sts:*
   23: em_sy:*
   24: em_un:*
   25: em_vi:*
   26: gb_htg1:*
   27: gb_htg2:*
   28: gb_in1:*
   29: gb_in2:*
   30: gb_bal:*
   31: em_ba1:*
   32: em_hum3:*
   33: em_hum4:*
   34: gb_pr4:*
   35: gb_htg3:*
   36: gb_htg4:*
   37: gb_htg5:*
   38: gb_htg6:*
   39: gb_htg7:*
   40: em_htg1:*
   41: em_htg2:*
   42: em_htg3:*
   43: em_hum5:*

   44: gb_pl3:*
   45: gb_pr5:*
   46: gb_htg8:*
   47: gb_htg9:*
   48: gb_htg10:*
   49: gb_htg11:*
   50: gb_htg12:*
   51: gb_htg13:*
   52: gb_htg14:*
   53: gb_in3:*
   54: gb_htg15:*
   55: gb_htg16:*
   56: gb_htg17:*
   57: em_htg4:*
   58: em_htg5:*
   59: em_htg6:*
   60: em_htg7:*
   61: em_hum6:*
   62: gb_htg18:*
   63: gb_htg19:*
   64: gb_ba3:*
   65: em_htg8:*
   66: em_htg9:*
   67: em_htg10:*
   68: gb_pr6:*
   69: gb_pr7:*
   70: gb_htg20:*
   71: gb_htg21:*
   72: gb_htg22:*
   73: gb_htg23:*
   74: gb_ro:*
   75: gb_sts1:*
   76: gb_sts2:*
   77: gb_sy:*
   78: gb_un:*
   79: gb_vi1:*
   80: gb_vi2:*
   81: gb_pat1:*
   82: gb_pat2:*
   83: em_htg0:*
   84: gb_htg24:*
   85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3967	99.6	4265	74	AF050183	AF050183 Rattus no
2	3936.2	98.9	4140	74	AF058790	AF058790 Rattus no
3	3828.4	96.2	4539	74	AF058789	AF058789 Rattus no
4	3601.4	90.5	4063	74	AF048976	AF048976 Rattus no
5	3579.6	89.9	4801	74	AB016962	AB016962 Rattus no
6	920.8	23.1	94770	69	HSDJ570F3	AL050732 Human DNA
c 7	920.8	23.1	150956	70	AL161903	AL161903 Homo sapi
c 8	920.8	23.1	169323	70	AL161906	AL161906 Homo sapi
9	629.4	15.8	4368	10	AF047711	AF047711 Homo sapi
10	443.2	11.1	829	74	AF053938	AF053938 Rattus no
11	336.4	8.5	4287	34	AK024488	AK024488 Homo sapi
c 12	240.2	6.0	56366	39	AC019800	AC019800 Drosophil
13	240.2	6.0	191504	36	AC012162	AC012162 Drosophil
c 14	240.2	6.0	202741	36	AC012161	AC012161 Drosophil
15	240.2	6.0	300994	28	AE003506	AE003506 Drosophil
16	235.2	5.9	3984	28	AB011280	AB011280 Caenorhab
17	222	5.6	65961	38	AC016523	AC016523 Drosophil
18	206.8	5.2	2385	34	AK022662	AK022662 Homo sapi
c 19	196	4.9	17027	72	AL365274	AL365274 Homo sapi
20	194.4	4.9	211	74	AF055883	AF055883 Rattus no
c 21	187.6	4.7	169323	70	AL161906	AL161906 Homo sapi

Chen

ALVHILQSTGKAKDFLSDMAMSEVDRFMEREHLIFRENTLAFKAIEEYMRLIGQKYLK
DAIGEFIRALYESEENCEVDPIKCTASSLAEHQANLRMCCLLAICKVVNSHCVFPREL
KEVFASWRLRCAERGREDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTL
TLIAKVIONLANFSKFTSKEDFLGFMNEFLELEWGSMOQFLYFISNLDTLTNSSSFEG
YIDLGRELSTLHALLWEVLPQLSKEALLKIGPLPRLLSDISTALRNPNIQRQPRQSE
RARSQPMVLRGPSAEMQGYMMRDLNSSIDLQSFMARGLNSSMDMARLPSPTKEKPPPP
PPGGGKDLFYVSRPPLARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGPG3GRL
NSSSVSNLAAVGDLLHSSQASLTAALGLRPAPAGRLSOGSGSSITAAGMRLSQMGVTT
DGVPAQOLRIPLSFQNPLFHMAADGPGPPAGHGGSSGHGPPSSHHHHHHHRRGGEP
PGDTFAPFHGYSKSEDLSTGVPKPPAASLHSHYSDEFGPSGTDFTRRQLSLQDNLQ
HMLSPPQITIGPQRPAPSGPGGSGGSGGSGGGGGGGQPPLQRGKSQOLLTVSAAQKPRPS
SGNLLQSPEPSYGPARPRQQSLSKEGSIGGSGGSGGGGGLKPSITKQHSQTPSTLN
PTMPASERTVAWVSNMPHLSADIESAHIEREEYKLKEYSKSMDESRLDRVKEYEEEIH
SLKERLHMSNRKLEEYERRLLSQEEQTSKILMOYQARLEQSEKRLRQQOVEKDSQIKS
IIGRLMLVEEELRRDHPAMAEPLPEPKKRLLDAQVERQLPPLGPTNPRVTLAPPWNGL
APPAPPPPPRLQITENGEFRNTADH"

BASE COUNT   936 a   1358 c   1203 g   768 t
ORIGIN

Best Local Similarity   99.6%;   Score 3967;   DB 74;   Length 4265;
Query Match 99.9%;   Pred. No. 0;
Matches 3981; Conservative   0; Mismatches   0; Indels   4; Gaps   1;

```
Qy      1 tagggagagactgagctgcccaagcacccatttccatttgtccccagcaaggcttc 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 TAGGGAGAGACTGAGCTGCCCCAAGCACCCATTTCCATTTGCTCCCCAGCAAGGCTTC 153

Qy     61 ctgagccggaggtaaaagctccatcaaagctacaaagtcacaacccaaacttgaccgg 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 CTGAGCCGGAGGCTAAAAGCTCCATCAAAGTCACAAAGTCACAACCCAAACTTGACCGG 213

Qy    121 accagcagcttcgacagatcctgcttccgaagtgctgaccatgacggcccgg 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 ACCAGCAGCTTCGACAGATCCTGCCTCGCCTTCCGAAGTGCTGACCATGACCGGGCCCGG 273

Qy    181 ctgatgcagagctcaaggagtctcactcccatgagtccctgcgactcccagcagtgct 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CTGATGCAGAGCTTCAAGGAGTCTCAAGGAGTCTCACTGCCCTGCGAGTCCCAGCAGTGCT 333

Qy    241 gctgaggccctgagctcaacctggatgaagactccattatcaagccagtacacagctcc 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GCTGAGGCCCTGGAGCTCCAACCTGGATGAAGACTCCATTATCAAGCCAGTACACAGCTCC 393

Qy    301 atcctgggccaggagttctgctttgaggtaacaacatcgtctggacaaaatgttttgcc 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 ATCCTGGGCCAGGAGTTCTGCTTTGAGGTAACAACATCGTCTGGGACAAAATGTTTTGCC 453

Qy    361 tgtcggtctgcagccgaaaggacaaatggattgagaatctacagagggctgtgaaaccc 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 TGTCGGTCTGCAGCCGAAAGGACAAATGGATTGAGAATCTACAGAGGGCTGTGAAACCC 513

Qy    421 aacaaggacaacagcgccgggtagataacgtgctgaaactatggatcatagaagctcga 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 AACAAGGACAACAGCGCCGGGTAGATAACGTGCTGAAACTATGGATCATAGAAGCTCGA 573

Qy    481 gagctgccccaagaagcgatattactgcgagttgccggacgacatgtctctatgca 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GAGCTGCCCCAAGAAGCGATATTACTGCGAGTTATGCCTGGACGACATGTCTCTATGCA 633

Qy    541 cggaccacttcaagccccgcctcagcctcaggagacactgtctttgggcgagcacttc 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    634 CGGACCACTTCCAAGCCCCGCCTCAGCCTCAGGAGACACTGTCTTTGGGGCGAGCACTTC 693

Qy    601 gagtttaacaacctgcctgctcggcgctgcggctgcacctgtgatctgtaccgtgactcggac 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    694 GAGTTTAACAACCTGCCTGCTCGGCGCTGCGGCTGCGCTGCATCTGTACCGTGACTCGGAC 753

Qy    661 aaaaagcggaggaaggacaaggcaggctacgttggctcgcttccagtggccacc 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 AAAAAGCGGAGGAAGGACAAGGCAGGCTACGTTGGCTCGCTGTTCCAGTGGCCACC 813

Qy    721 ctggctgggcgccacttcacagagcagtggtaccccgtgaccctgacccagcaggaagtggg 780
```

ALIGNMENTS

RESULT 1
AF050183
LOCUS   AF050183   4265 bp   mRNA   ROD   15-SEP-2000
DEFINITION Rattus norvegicus GTPase activating protein SynGAP-c mRNA, complete
            cds.
ACCESSION   AF050183
VERSION   AF050183.2   GI:10140854
KEYWORDS   .
SOURCE   Norway rat.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 4265)
  AUTHORS   Kim,J.H., Liao,D., Lau,L.F. and Huganir,R.L.
  TITLE   SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
            protein family
  JOURNAL   Neuron 20 (4), 683-691 (1998)
  MEDLINE   98240917
REFERENCE   2 (bases 1 to 4265)
  AUTHORS   Kim,J.H. and Huganir,R.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-FEB-1998) Neuroscience, Johns Hopkins Medical
            Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
COMMENT   On Sep 15, 2000 this sequence version replaced gi:3044054.
FEATURES   Location/Qualifiers
     source   1..4265
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="hippocampus"
                     /dev_stage="adult"
     CDS   277..3792
                     /note="neuronal synaptic RasGAP; associates with the
                     PSD-95/SAP90 protein family; alternatively spliced
                     variants, SynGAP-a and SynGAP-b"
                     /codon_start=1
                     /product="GTPase activating protein SynGAP-c"
                     /protein_id="AAC40082.2"
                     /db_xref="GI:10140855"
                     /translation="MQSFKESHSHESLLSPSSAAEALELNLDEDSIIKPVHSSILGQE
                     FCFEVTTSSGTKCFACRSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELP
                     PKKRYCELCLDDMLYARTTSKPRSASGDTVFWGEHFEFENNLPAVRALRLHLYRDSDK
                     KRKKDKAGYVGLVTVPVATLAGRHFTEQWYPVTLPTGSSGGMGSGGGGGSGGSGGSGG
                     KGKGGCPAVRLKARYQTMSILPMELYKEFAEYVTNHYRMLCAVLEPALNVKGKEEVAS
```

c 22  157     3.9   81971     HS593C16      AL035702 Human DNA
c 23  157     3.9  156700  50  AC027051      AC027051 Homo sapi
  24  139.2   3.5  155332  50  AC040900      AC040900 Homo sapi
c 25  139.2   3.5  156700  50  AC027051      AC027051 Homo sapi
c 26  113.4   2.8    4107  74  MMU20238      U20238 Mus musculu
  27  103.4   2.6  159777  35  AC011703      AC011703 Drosophil
  28  103.2   2.6  157803  35  AC011492      AC011492 Homo sapi
  29   99.2   2.5    2837  35  HSINSP4BP     X89399 Homo sapien
  30   98.4   2.5   39660  69  CEC07B5       Z46266 Caenorhabdi
  31   97.6   2.5  237619  56  AC073765      AC073765 Mus muscu
  32   98.4   2.5     833  28  AB011285      AB011285 Caenorhab
  33   95.8   2.4    2647  81  E13125        E13125 Bovine gene
  34   95.8   2.4    2730   3  BTU30857      U30857 Bos taurus
  35   93.8   2.4     760  28  AB011283      AB011283 Caenorhab
  36   93.8   2.4    2276  28  AB011279      AB011279 Caenorhab
c 37   89.2   2.2     314  75  G21290        G21290 human STS W
c 38   71.2   1.8     310  28  AB011284      AB011284 Caenorhab
c 39   70.8   1.8    5719  74  MMCATS1       AF051726 Mus muscu
c 40   70.8   1.8  180385   9  AC007461      AC007461 Homo sapi
c 41   69.6   1.7    3720  80  S76368        S76368 ORF 5' of E
c 42   69.6   1.7   43658  80  HSV3PRGEN     M86409 Herpesvirus
c 43   69.6   1.7  112930  80  HSGEND        X64346 Herpesvirus
c 44     68   1.7    5579   8  AB011110      AB011110 Homo sapi
c 45   67.4   1.7  314556  56  AC073495      AC073495 Mus muscu

```
Db   814  |||||||||||||||||||||||||||||||||||||                              CTGGCTGGCGCCACTTCACAGACAGTAGCAGTGG  873
Qy   781  ggctctggggtatggctcggggggaggggtcaggggcggctcaggggcaa                                                           840
Db   874  |||||||||||||||||||||||||||||||||||                                GGCTGCTGGCGGTATGGCTCGGGGGCTGGGGGAGGGAGGGGTCAGGGGGCAA  933
Qy   841  gggaaaggaggctgtcctgctcgtgcggtaccagacaatgagtatcctg                                                            900
Db   934  |||||||||||||||||||||||||||||||                                   GGGAAAGGAGGCGTGTCCTGCTCGTGCGGTACCAGACCGGTTACCACACAATGAGTATCCTG  993
Qy   901  cccatggagctataaggagttcttgcgaatatgtgaccaaccactaccgcatgctgtgt                                                  960
Db   994  |||||||||||||||||||||||||                                         CCCATGGAGCTATPATAAGGAGTTTGCAGAATATGTGACCAACCACTACCGCATGCTGTGT 1053
Qy   961  gccgtgctggagcccgcctcaatgtcaaggcaaggaggagtcgtagtgacctggtt                                                     1020
Db  1054  ||||||||||||||||||||||||||||                                      GCCGTGCTGGAGCCGCCGCCTCAATGTCAAGGCAAGGAGGGCGGTCGCTAGTGCACTGGTT 1113
Qy  1021  cacatcctgcaaagcacagacaagccaaggaccttctcttcagacatgatgtcagag                                                    1080
Db  1114  |||||||||||||||||||||||||||||||                                   CACATCCTGCAAAGCACAGCAAGGCAAGGCCAAGGACTTCCTTTCAGACACATGGCCAATGTCAGAG 1173
Qy  1081  gtagaccggttcatggagcgggaacacactcatattccgcgagaacacgctcgccactaa                                                1140
Db  1174  |||||||||||||||||||||||||||||||                                   GTAGACCGGTTCATGGAGCGGGAACACACTCATATTCCGCGAGAACACGCTCGCCCACTAAA 1233
Qy  1141  gccatagagagtatatgagactgatggccagaaatacctcaaggatgccattggggag                                                  1200
Db  1234  |||||||||||||||||||||||||||||||                                   GCCATAGAGAGAGTATATGAGACTGATGGCCAGAAATACCTCAAGGATGCCATTGGGGAG 1293
Qy  1201  ttcatccggggtctgtatgaatctgaaggagaactgctgaagtagacccatcaagtgcaca                                               1260
Db  1294  ||||||||||||||||||||||||||||||                                    TTCATCCGGGGTCTGTATGAATCTGAGGAGAACTGTGAAGTAGACCCCATCAAGTGCACA 1353
Qy  1261  gcgtccagtctggcagagacccaggccaacctgcggatgtctgtagttggccctgtgc                                                   1320
Db  1354  ||||||||||||||||||||||||||||||                                    GCGTCCAGTCTGGCAGAGACCCAGGCCAACCTGCGGATGTGCTGTGAGTTGGCCCTGTGC 1413
Qy  1321  aaggtggtcaactccattgcgtgttcccgagggagctgaaaggaggtgtttgatcatcatgg                                              1380
Db  1414  |||||||||||||||||||||||||||||                                     AAGGTGGTCAACTCCCATTGCGTGTTCCCGAGGGAGCTGAAAGGAGGTGTTTGCATCATCGG 1473
Qy  1381  cggctgcgtgtcagagcggggccgggaggacattgctgacaggctgatcagccgcctcg                                                  1440
Db  1474  |||||||||||||||||||||||||||||                                     CGGCTGCGCTGTGCGGAGAGCGGGGGCCGGGACAATTGCTGACAGGCTGATCAGCGCCCTCG 1533
Qy  1441  ctcttcctgcgcttcctctgcccggccatcaatgtgccagtctgtttggactgatgcag                                                  1500
Db  1534  ||||||||||||||||||||||||||||||                                    CTCTTCCTGCGCTTCCTCTGCCCGGCCATCATGTGCCCATCATGTGCTGTTTGGACTGATGCAG 1593
Qy  1501  gagtaccagatagcagacctcacgagtcaattcctgtatagagatatccaaggttatccagaac                                            1560
Db  1594  ||||||||||||||||||||||||||||||                                    GAGTACCAGATGAGCAGACAGACCTCACGAGTCAATTCTTGTATGCAGATATTCCAACCTGACACA 1653
Qy  1561  ctggccaacttttccaagtttacctcaaagagggacttcctggctt catgaacagttt                                                  1620
Db  1654  ||||||||||||||||||||||||||||||                                    CTGGCCAACTTTTCCAAGTTTACCTCAAAGGAGGACTTCCTGGCTTCATGAACGCAGTTT 1713
Qy  1621  ctggagctggaatgggttctatgcagcaattcttgtatgagatatccaacctggacaca                                                  1680
Db  1714  ||||||||||||||||||||||||||||||                                    CTGGAGCTGGAATGGGGTTCTATGCAGCAATTCTTGTATGCAGATATTCCAATCGGACACA 1773
Qy  1681  ctgaccaacaacagagcagtgtttgagggctacatagactggtgggccgcgagctcccacactt                                            1740
Db  1774  ||||||||||||||||||||||||||||||                                    CTGACCAACAACAGCAGCAGTTTTGAGGGCTACATAGACTGGTGGGCCCCGAGCTCTCCACACTT 1833
Qy  1741  cacgcccctgctctggggagtgctgcccccagctcagcaaggagccctcctgaagctgggc                                                1800
Db  1834  ||||||||||||||||||||||||||||||                                    CACGCCCCTGCTCTGGGAGGTGCTGCCGGAGTGCCCCCAGCTCAGCAAGGAAGCCCTGAAAGCTGGGC 1893
Qy  1801  ccgctgcccggctccctcagcgacatcagcacagccctgaggaacctaacatccaaagg                                                  1860
```

```
Db  1894  |||||||||||||||||||||||||||||||||||                               CCGCTGCCCCGGCTCCTCAGCGACATCAGCACAGCCCTGAGGACCCTAACATCCAAGG 1953
Qy  1861  cagccagcagccgccagagcgagcgcgctcggtctcagccacatggtgctgcgcggggccgtca                                             1920
Db  1954  ||||||||||||||||||||||||||||||                                    CAGCCCAGCCCCAGAGCGAGCGGCGCGGCTCGGTCTCAGCCCATGGTGCTGCCGGGCCGTCA 2013
Qy  1921  gccgaaatgcaggctacatgcggcgacctcaacagctccatcgacctcagtcctic                                                     1980
Db  2014  |||||||||||||||||||||||||||||||                                   GCCGAAATGCAGGGCTACATGATGCGGCGACCTTCAACAGCTCCATCGACCTTCAGTCCTTC 2073
Qy  1981  atggctcgaggcctcaacagctctatggacatggctgcctcccctcccccaaccaaggag                                                 2040
Db  2074  ||||||||||||||||||||||||||||||                                    ATGGCTCGAGGCCTCAACAGCTCTATGGACATGGCTGCTCCCTCCCCCAACCAAGGAG 2133
Qy  2041  aaaccccgcgccccctccggtggggtaaaagaccgttctatgtgagccggccacca                                                     2100
Db  2134  ||||||||||||||||||||||||||||||                                    AAACCCCGCGCCCCCTCCGGTGGGGTAAAGACCTGTTCTATGTGAGCCGGCCACCA 2193
Qy  2101  ctggccggtcctcccagcatactgcacgagcagctcggacatcacagagccggagcag                                                   2160
Db  2194  ||||||||||||||||||||||||||||||                                    CTGGCCCGGTCCTCCCCAGCATACTGCACGAGCAGCTCGGACATCACAGAGCCGGAGCAG 2253
Qy  2161  aagatgctgagtgtcaacaagagtgtgtccatgctgaccctgcaggggcgacgggcctggg                                                2220
Db  2254  |||||||||||||||||||||||||||||                                     AAGATGCTGAGTGTCAACAAGAGTGTGTCCATGCTGACCCTGCAGGGCCGACGGGCCTGGG 2313
Qy  2221  ggccgccttaacagcagtagtgttgttccaacctgcagctgttgggacctgttgcactca                                                 2280
Db  2314  ||||||||||||||||||||||||||||||                                    GGCCGCCTTAACAGCAGTAGTGTTCCAACCTGCAGCTGTGTTGGGACCTGTGTTGCACTCA 2373
Qy  2281  agccaggcttcactgacacgagccttggggttcggcctgcacctgccgggcgcctctcc                                                  2340
Db  2374  |||||||||||||||||||||||||||||                                     AGCCAGGCTTCACTGACACGAGCCTTGGGGTTCGGCCTGCACCTGCCGGGCGCCCTCTCC 2433
Qy  2341  caaggagtggctgctcttccatcacagcagccggatgccgatgctgccagatgggtgtcact                                               2400
Db  2434  |||||||||||||||||||||||||||||                                     CAAGGAGTGGCTGCTCTTCCATCACCACATCACCACCGGATGCCGATGCTGCCAGATGGGTGTCACT 2493
Qy  2401  acggatggtgtcccccgcccagcaactgcgcatccctcttccttccagaaccctctctc                                                  2460
Db  2494  |||||||||||||||||||||||||||||                                     ACGGATGGTGTCCCCCGCCCCAGCAACTGCGCATCCCTCTTCCTTCCTTCCAGAAACCCTCTCTTC 2553
Qy  2461  catatggctcgccgatggacaggccccgcccagcaggccatggaggacgcagtggccatggt                                               2520
Db  2554  |||||||||||||||||||||||||||||                                     CATATGGCTCGCCGATGGACCAGGCCCCGCCCCAGCAGGCCATGGAGGACGCAGTGGCCATGGT 2613
Qy  2521  ccacctccttcctccatcaccaccaccaccaccaccaccaccgaggggagaaccccca                                                   2580
Db  2614  |||||||||||||||||||||||||||||                                     CCACCTTCCTCCATCACCACCATCACCACCATCACCACCGAGGGGGAGGGGGAGAACCCCCA 2673
Qy  2581  ggggacactttgcccgttccatggctatagaggcgaggacctctcacgggtc                                                         2640
Db  2674  |||||||||||||||||||||||||||||                                     GGGGACACTTTTGCCCCGTTCCATGGCTATAGCAGCAGGAGGACCCTCTCACAGGGGTC 2733
Qy  2641  cctaagcccctgcggcctccatccttcacagccacagtcagtgatgagtttggaccc                                                    2700
Db  2734  |||||||||||||||||||||||||||||                                     CCTAAGCCCCCTGCGGCCTCCATCCTTCACCACGCCACAGTCAGTGATGAGTTTGGACCC 2793
Qy  2701  tctggtactgattttaccggtgccgagctctcacttcaggacaacctacagcacatgctc                                                 2760
Db  2794  |||||||||||||||||||||||||||||                                     TCTGGTACTGATTTTTACCCGGTGCCGAGCTCTCCACTTCAGGACAACCTACAGCACATGCTC 2853
Qy  2761  tccccgcccagatcaccatcggtcccagtccccaggggtccccagggggggc                                                          2820
Db  2854  |||||||||||||||||||||||||||||                                     TCCCCGCCCCAGATCACCATCGGTCCCCAGTCCCCAGGGGCCCCAGGGGGGC 2913
Qy  2821  agtggtggggggcagtggtggtgggggcggtgggggcgtggggggcagccagccctgcaaa                                                2880
Db  2914  |||||||||||||||||||||||||||||                                     AGTGGTGGGGGCAGTGGTGGGGGGCGGTGGGGGGCGTGGGGGGCCAGCCACCTGCAAA 2973
Qy  2881  tctcagcagttgacagtgagtgctgccagaaacccggccgtccagcgggaacctattg                                                   2940
Db  2974  |||||||||||||||||||||||||||||                                     TCTCAGCAGTTGACAGTGAGTGCTGCCCAGAAACCCGGCCGTCCAGCGGGGAACCTATTG 3033
```

```
RESULT      2
AF058790
LOCUS       AF058790        4140 bp    mRNA            ROD       09-OCT-1998
DEFINITION  Rattus norvegicus SynGAP-b mRNA, complete cds.
ACCESSION   AF058790
VERSION     AF058790.1  GI:3722228
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4140)
  AUTHORS   Kim,J.H., Liao,D., Lau,L.F. and Huganir,R.L.
  TITLE     SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
            protein family
  JOURNAL   Neuron 20 (4), 683-691 (1998)
  MEDLINE   98240917
REFERENCE   2  (bases 1 to 4140)
  AUTHORS   Kim,J.H. and Huganir,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-APR-1998) Neuroscience, Johns Hopkins Medical
            Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
REFERENCE   3  (bases 1 to 4140)
  AUTHORS   Kim,J.H. and Huganir,R.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-1998) Neuroscience, Johns Hopkins Medical
            Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
REMARK      Sequence update by submitter
COMMENT     On Oct 9, 1998 this sequence version replaced gi:3065890.
FEATURES             Location/Qualifiers
     source          1..4140
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /tissue_type="hippocampus"
     CDS             1..3750
                     /note="neuronal RasGAP; contains PH domain, C2 domain, GAP
                     domain in the N-terminal region, and a T/SXV motif at the
                     C-terminus"
                     /codon_start=1
                     /product="SynGAP-b"
                     /protein_id="AAC63511.1"
                     /db_xref="GI:3722229"
                     /translation="MGLRPPTPTPSGGSGSGSLPPPSHRQPLRRCSSCCFPGEYHLG
                     RSRRKSVPGGKQYSMEAAPAAPPRPSOGFLSRRLKSSIKRTKSQPKLDRTSSFRQlLP
                     RFRSADHDRARLMQSFKESHSHESLLSPSSAAEALELNLDEDSIIKPVHSSILGQEεC
                     FEVTTSSGTKCFACRSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELPPK
                     KRYCELCLDDMLYARTTSKPRSASGDfVFWGEHFEFNNLPAVRALRLHlYRDSDKKR
                     KKDKAGYVGLVTVPVATLAGRHFTEQWYPVTLPTGSGSGGMGSGGGGSGGGSGGGKG
                     KGGCPAVRLKARYQTMSLlPMELYKEFAEYVTNHYRMLCAVLEPALNVKGKEEVASAL
                     VHILQSTGKADFLSDMAMSEVDRFMEREHLIFRENTLATKAIEEYMRLIGQKYLKDA
                     IGEFIRALYESEENCEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKE
                     VFASWRLRCAERGREDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTLTL
                     IAKVIQNLANFSKFTSKEDFLGFMNEFLELEWGSMQQFLYEISNLDTLTNSSSFEGYI
                     DLGRELSTLHALLWEVLPQLSKEALLKLGPLPRLLSDISTALRNPNIQRQPSRQSERA
                     RSQPMVLRGPSAEMQGYMWRDLNSSIDLQSFMARGLNSSMDMARLPSPTREKPPPPP
                     GGGKDLFYVSRPPLARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGDPGGRLNS
                     SSVSNLAAVGDLLHSSQASLTRAALGLRPAPAGRLSOGSGSSITAAGMRLSQMGVTTDG
                     VPAQQLRIPLSFQNPLFHMAADCGPPAGHGSSSGHGPPSSHHHHHHHRRGGEPPG
                     DTFAPFHGYSKSEDLSTGVPKPPAASIIHSHSYSDEFGPSGTFTRRQLSLQDNLQHM
                     LSPPQITIGPQRPAPSGPGGSGSGGSGGGGGGGQPPLQRGKSQQLTVSAAQKPRPSSG
                     NLlQSPEPSYGPARPRQQSLSKEGSICGSGGSIGGSGGGGGLKPSITKQHSQTPSTLNPT
                     MPASERTVAWVSNMPHLSADIESAHIEREEYKLKEYSKSMDESRLDRVKEYEEEIHSL
                     KERlHMSNRKLEEYERRLLSQEEQTSKILMQYQARLEQSEKRLRQQVERDSQIKSII
                     GRLMLVEEELRRDHPAMAEPLPEPKKRLLDAQRGSFPPWVQQTRV"
BASE COUNT        902 a       1328 c       1182 g        728 t
ORIGIN
```

```
Qy   2941 cagtccccggaaccaagttatggtcctgccgtccacgcaacagagcctcagcaaagag 3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3034 CAGTCCCCGGAACCAAGTTATGGTCCTGCCCCGTCCACGCAACAGAGCCTCAGCAAAGAG 3093

Qy   3001 ggcagcattggggcagggggcagcggtggcggtggccagccagggggtggggggctcaagccctcc 3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3094 GGCAGCCATTGGGGCAGCGGCGGCGGCAGCCGCGTGGCGGAGGGGTGGGGGGCTCAAGCCCTCC 3153

Qy   3061 atcaccaagcagcattcccagactccatccacgctgaacccacgatgccggcctcggag 3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3154 ATCACCAAGCAGCATTCCCAGACTCCATCCACGCTGAACCCACGATCCGGCCTCCGGAG 3213

Qy   3121 cggactgtagcctggttgtccaatatgcctcacctgtccgctgacatcgagagtgcacac 3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3214 CGGACTGTAGCCTGGTGTCCAATATGCCTCACCTGTCCGCTGACATCGAGAGTGCACAC 3273

Qy   3181 attggacgggaagagtacaagctgaaggagtactccgaagtccatggacgagagccgactg 3240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3274 ATTGGACGGGAAGAGTACAAGCTGAAGGAGTACTCCGAAGTCCATGGACGAGAGCCGACTG 3333

Qy   3241 gacagggtgaaggagtacgagaggagatcccactcactgaaggaaggctacacacatgtcc 3300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3334 GACAGGGTGAAGGAGTACGAGGAGGAGATCCCACTCACTGAAGGAAAGGCTACACACATGTCC 3393

Qy   3301 aaccggaagctggaagagtacgagcggaggctgctgtccaggaagagcagaccagcaag 3360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3394 AACCGGAAGTCTGGAAGAGTACGAGCGGAGGCTGCTGTCCCAGGAAGAGCCAGACCAGCAAG 3453

Qy   3361 atcctgatgcagtaccaaagcccgcctgagcagagcgagaagcgcttgaggcagcagcag 3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3454 ATCCTGATGCAGTACCAAAGCCCGCCTGGAGCAGAGCAGCCGCTTGAGGCAGCAGCAG 3513

Qy   3421 gtggagaaggactccagatacaagagacatcattggcaggctgatgctggtggtggaggag 3480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3514 GTGGAGAAGGACTCCAGATACAAGAGCATCATTGGCAGGCTGATGCTGGTGGTGGAGGAGGAG 3573

Qy   3481 ctgcgcgggaccaccccgccatgctgagccgcgtgcctgaacccaagagaagctgtctc 3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3574 CTGCGCGGGACCACCCCGCCACCCCGCCATGCTGAGCCGCTGCCTGAACCCAAGAGAGGCTGCTC 3633

Qy   3541 gacgctca---gagaggcagcttccccccttgggtccaacaaacccgcgtgtgacgctg 3596
          ||||||||   |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3634 GACGCTCAGGTGGAGAGGCAGCTTCCCCCCTTGGGTCCAACAAACCCGCGTGTGACGCTG 3693

Qy   3597 gccccaccttggaacggcctggccccccacccccaccccgcggctgcgagatc 3656
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3694 GCCCCACCTTGGAACGGCCTGGCCCCCTGGCCCCCACCCCGCCACCCCGGCTGCAGATC 3753

Qy   3657 acagagaacggcgagttccggaacaccgcagacaccagccaccactagccagcatcacagacct 3716
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3754 ACAGAGAACGGCGAGTTCCGGAACACCGCAGACACCGGCCACCAGCCATCACAGACCT 3813

Qy   3717 ccttcctgcaccctacccggccccaccgggtcacagacctccttcccagtgcacc 3776
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3814 CCTTCCCTGTGCACCTACCCCGGCCACCCGGCGTCACAGACCTCCTTCCCAGTGCACC 3873

Qy   3777 cgaccctggaacactcaccaaccaccaggactggacgtcaccaagggacagcgggattgtc 3836
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3874 CGACCCTGGAACACATCACCACCAGACCAGGAAGGGACACTGGGACACCAGCCGGATTGTC 3933

Qy   3837 tcccttaacgcctccttggggacgcaccccatcgtgtcaacccccactgtctcattccaggagggga 3896
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3934 TCCCTTAACGCCTCCTTGGGGACGCACCCCATCGTCAACCCCACTGTCTCCATTCCAGGAGGGA 3993

Qy   3897 gagtggagaccctcagtcagctgcctctcacccccagacagacaccacccacacagacaccccctt 3956
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3994 GAGTGGGACCCCTCAGCTGCCTCCTCCAGACACCACCCACCACCACCAGACACCCCCTT 4053

Qy   3957 cactctggggtgctatccccatcct 3981
          |||||||||||||||||||||
Db   4054 CACTCTGGGGTGCTATCCCCATCCT 4078
```

```
Query Match             98.9%;  Score 3936.2;  DB 74;  Length 4140;
Best Local Similarity   99.7%;  Pred. No. 0;
Matches 3944;  Mismatches    0;  Conservative    0;  Indels   13;  Gaps   0;
```

```
Qy    25  gcacccccattcccattgctcccagcaaggcttcctgagccggagggctaaaaagctcc    84
          || ||||||||||||||||||||||||||||||
Db   184  GCCCCGCTGGCGCCCTTCCGCGCCTTCGCAAGGCTTAAAAAGCTCC                243

Qy    85  atcaaacgtacaaagtcacaaccaaagttgacacggaccagcagcttcgacagatcctg  144
          |||||||||||||||||||||||||||||||||||
Db   244  ATCAAACGTACAAAGTCACAACCCAAACTTGACCGGACCAGCAGCTTTCGACAGATCCTG  303

Qy   145  cctcgcttccgaagtgctgacacggggccccggctgatgcagagcttcaaggagtct    204
          |||||||||||||||||||||||||||||||||
Db   304  CCTCGCTTCCGAAGTGCTGACCATGACCGGGGCCCGGCTGATGCGGAGAGCACCAG     363

Qy   205  cactccatgagtccctggtcagtccaacagtgctgctgagcctgagctcaacctg     264
          |||||||||||||||||||||||||||||||
Db   364  CACTCCCATGAGTCCCTGCTGAGTCCAGCAGTCCTGCTGAGGCCCTGAGCTCAACCTG   423

Qy   265  gatgaagactccattatcagcaagtacacagctccatcctgggccaggagttctgcttt  324
          ||||||||||||||||||||||||||||||
Db   424  GATGAAGACTCCATTATCAGCAAGTACACACAGCTCCATCCTGGGCCAGGAGTTCTGCTTT  483

Qy   325  gagtaacaacatcgtctgggacaaaaatgtttgcctgtcggtctgcagccgaaaaggac  384
          |||||||||||||||||||||||||||||||
Db   484  GAGGTAACAACATCGTCTGGGACAAAAATGTTTGCCTGTCGGTCTGCAGCCGAAAAGGAC  543

Qy   385  aaatggattgagaatctacagagggctgtgaaacccaacaaggacaacgcgccgggta   444
          |||||||||||||||||||||||||||||||
Db   544  AAAATGGATTGAGAATCTACAGAGGGCTGTGAAACCCAACAAGGACAACGCCCGGGGTA   603

Qy   445  gataacgtgctgaaactatggatcatagaagctcgagagctgcccccccagagagcgatat  504
          |||||||||||||||||||||||||||||||||
Db   604  GATAACGTGCTGAAACTATGGAATCATAGAAGCTCGAGAGCTGCCCCCCCAAGAGAGCGATAT  663

Qy   505  tactgcgagttatgcctggactgacgacatgctctatgcacggaccacttccaagccgctca  564
          |||||||||||||||||||||||||||||||
Db   664  TACTGCGAGTTATGCCTGGACTGACTGACGACATGCTCTATGCACGGACCACTTCCAAGCCGCTCA  723

Qy   565  gcctcaggagacactgtctttttgggcgcagcacttcgagttaacaacctgctgctgtc  624
          |||||||||||||||||||||||||||||||
Db   724  GCCTCAGGAGACACTGTCTTTTGGGGCGCAGCACTTCGAGTTTAACAACCTGCTGCTGTC  783

Qy   625  cgggcgctgcggctgcatctgaccgtgactcggacaaaaagcggaagaagaagaagacaagcca  684
          |||||||||||||||||||||||||||||||||
Db   784  CGGGCGCTGCGGCTGCATCTGTACCGTGACTCGGACAAAAAGCGGAAGAAGGACAAGGCA  843

Qy   685  ggctacggttggcctgactgttccagtggccacctgggcgccacttcacagag        744
          |||||||||||||||||||||||||||||||
Db   844  GGCTACGTTGGCCTGACTGTTCCAGTGGCCACCTGGGCCACTTCACAGAG          903

Qy   745  cagtggtaccccgtgacctgcaaacaggaagtggggggctctggggtatgggctcggggg   804
          |||||||||||||||||||||||||||||||
Db   904  CAGTGGTACCCCGTGACCCTGCAACAGGAAGTGGGGGCCTCTGGGGTATGGGCTCGGGG   963

Qy   805  gggaggggggtcaggggcggctcaggggcaaaggggaaggaggctgtcctgctgtg      864
          |||||||||||||||||||||||||||||||
Db   964  GGAGGGGGGTCAGGGGCGGCTCAGGGGCAAAGGGGAAAAGGGAAAGGAGGAGGCTGTCCTGCTGTG  1023

Qy   865  cggctgaaggcccgttaccagacagacaatagttatcctgccatggagttatataaaggagttt  924
          |||||||||||||||||||||||||||||||||
Db  1024  CGGCTGAAGGCCCCGTTACCAGGACCCTATCCTGCCCATGGAGCTATATAAGGAGTTT  1083

Qy   925  gcagaatagtgaccaaccactacccgctgctgtgtgccgtgctgtgctgtcacactggttcacatctgtgagc  984
          |||||||||||||||||||||||||||||||||
Db  1084  GCAGAAATATGTGACCAACCACTACCCGCTGTGTGCGGTTCACATCTGCAAAGCACAGGCAAG  1143

Qy   985  gtcaaggcaaggagagtcggctagtgcactggttcacatcctgcaaagcacaggcaag  1044
          |||||||||||||||||||||||||||||||
Db  1144  GTCAAGGCAAGGAGGAGTCGGCTAGTGCACTGGTTCACATCCTGCAAAGCACAGGCAAG  1203

Qy  1045  gccaagagactccttcagacatggccagaggtagaccggttcatggagcgggaa       1104
          |||||||||||||||||||||||||||||||
Db  1204  GCCAAGAGACTTCCTTCTTTCAGACATGGCCCATGTCAGAGGTAGACCGGTTCATGGAGCGGGAA  1263
```

```
Qy  1105  cacctccatattccgcgagaacacgctctcgccactaaagctccgagagagtatatgagactg  1164
          |||||||||||||||||||||||||||||
Db  1264  CACCTCCATATTCCCGAGAACACACGCTCGCCACTAAAGCCTAAAGCAGTATATGAGACTG  1323

Qy  1165  attggccagaaatacctcaaggagtccattgggaggttcatccggctctgtatgaatct  1224
          |||||||||||||||||||||||||||||
Db  1324  ATTGGCCAGAAATACCTCAAGGATGCCATTGGGAGGTTCATCCGGCTCTGTATGAATCT  1383

Qy  1225  gaggagaaactgtgaaggtagcaccaccatccagtgcacagggtccagttctggcagagcaccag  1284
          ||||||||||||||||||||||||||||||
Db  1384  GAGGAGAAACTGTGAAGTAGAGACCCCATCAAGTGCACAGGGTCCAGTGGCAGAGCACCAG  1443

Qy  1285  gccaacctcggagtgtcgtgattggccctgcagaggtggtcaactccattgctg        1344
          | ||||||||||||||||||||||||||||
Db  1444  GCCAACCTCGGATGTGCTGTGATGGGCCCCGTGTGCAAGGTGGTGTCAACTCCCATTGCTG  1503

Qy  1345  ttcccgagggagctgaagagggtgtttgcatcatgtggcgctggtctgcagacgggc     1404
          ||||||||||||||||||||||||||||||
Db  1504  TTCCCGAGGGAGCTGAAGGAGGTGTTTGCAATCATGTGGCCAACTTTCCAAGTTTACC  1563

Qy  1405  cgggaggacattgctgacaggctgatcagcgcctcgctcttcctgggttctgctgcccg  1464
          |||||||||||||||||||||||||||||||
Db  1564  CGGGAGGACATTGCTGACAGGCTGATCAGCGCCTCGCTCTTCCTCGCGCTTCCTGTGCCCG  1623

Qy  1465  gccatcatgtcgcccagtctgtttggactgatgatgagtagtaccagatgagcgagaccca  1524
          |||||||||||||||||||||||||||||||
Db  1624  GCCATCATGTCGCCCAGTTCGTTTGGACTGATGATGGCAGGAGTACCCAGATGAGCGAGAACCTCA  1683

Qy  1525  cgaaccctcaccctacatcgccaaggttatcagaacctggccaacttttccaagtttacc  1584
          ||||||||||||||||||||||||||||||||
Db  1684  CGAACCCTCACCCTCCACATCGCCCAAGGTTATCATCGACAACCTGGCCAACTTTTTCAAGTTTACC  1743

Qy  1585  tcaaaggagacttcctggcttcatgaacgagtttctggagctggaatggggttctatg  1644
          |||||||||||||||||||||||||||||||
Db  1744  TCAAAGAGGACACTTCCTGGCTTCATGAACGAGTTTCTGGAAGCGTTTCTGGAAGTGGGTTCTATG  1803

Qy  1645  cagcaattctgtgatgagatatccaacctgacccactgacaacagcagcagttttgag  1704
          |||||||||||||||||||||||||||||||
Db  1804  CAGCAATTCTTGTGATGAGATATCCAACCTGACACCTGACACGCAGCAGCAGTTTGAG  1863

Qy  1705  ggctacatagactggccgcgagctctccaacacttcacgcctgctctgggaggtgctg  1764
          |||||||||||||||||||||||||||||||
Db  1864  GGCTACATAGACTTGGGCCGCCGAGCTCTCCAACACTTCACGCCTGCTCTCGGGAGGTGCTG  1923

Qy  1765  cccagtcagcaggaagccctcctgaagtctggccccgctgccccggctctccagcgac  1824
          |||||||||||||||||||||||||||||||
Db  1924  CCCCAGCTCAGCAGGAAGCCCTCCTGAAGTCTGGCCCCGCTGCCCCGGCTCTCCAGCGAC  1983

Qy  1825  atcagcacagccctgaggaaccctaacctcaaagcgagcggagcccgcagggagcgc  1884
          |||||||||||||||||||||||||||||||
Db  1984  ATCAGCACCAGCCCTGAGGAACCCTAACATCCAAAGGCAGCAGCCCCAGAGCGGAGCGGC  2043

Qy  1885  gctcggtctcagcccatgtgctgcgcgggccgtcagccgaaatgcgggtcagagggtacatgatg  1944
          |||||||||||||||||||||||||||||||
Db  2044  GCTCGGTCTCAGCAGCCATGTGCTGCGCGGCGCCGGCCGGAAATGCGGGTCAGGGCGTACACTGATG  2103

Qy  1945  cgggacctgaagcgttacagagagcctagttcctcagtcctttcatggctcgaggcctcaacagctct  2004
          ||||||||||||||||||||||||||||||
Db  2104  CGGGACCTGCAACAGCTCCAGCTCCATTCAGTTCCTTCAGTGGCTCGAGGCCTCAACAGCTCT  2163

Qy  2005  atggacatggctctcgcctccccccctccccccccctccccggt                 2061
          |||||||||||||||||||||||||||||||
Db  2164  ATGGACATGGCTCTCGCCTCCCCCCCCCTCCCCAACCCCGCGCCCCTCCCGGT          2223

Qy  2065  ggggtaaagacgttctatcgtgacgtagaccggtcccaggcacaggcaag           2124
          |||||||||||||||||||||||||||||
Db  2224  GGGGTAAAGACGTTCTATGTGACGTAGGCGGCCGCCGTCCCCAGCATAC            2283

Qy  2125  tgcacgagcagctcagacatcagcagtggccggttcatggagtgtcagcaagagt      2184
          ||||||||||||||||||||||||||||||
Db  2284  TGCACGAGCAGCTCGACATCACAGAGCCGACATCCACAGAGCCGACATGAGTGTCAACAAGAGT  2343

Qy  2185  gtgtccatgctggacctcgagggcgacgggccctgggggcgccctcaacagcagtagtgtt  2244
```

Left column:

```
        ||||||||||||||||||||||||
Db 2344 GTGTCCATGCTGGACCTGCAGGCGACGGGCGACGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTGGGGGCCTTAACAGCAGTAGTGTT 2403
Qy 2245 tccaacctggcagctggcagcttgttgggacctgttcactcaagccaggcttcactgacacagcagcc 2304
        ||||||||||||||||||||||||||
Db 2404 TCCAACCTGGCCAGCTGTTGGGGACCTGTTGCGGGCCTCAAGCCAGGCTTCACTGACACGCAGCC 2463
Qy 2305 ttggggttgcggccctgcccggcgacctgcacctgcaggagttggctcttccatcaca 2364
        ||||||||||||||||||||||
Db 2464 TTGGGGTTGCGGCCCTGCCACCTGCAGGAGTGGCTCTTCCATCACA 2523
Qy 2365 gcagccggcatgcgcctcagccaggatgggtgtcactacggatggtgtccccgccagcaa 2424
        ||||||||||||||||||||||||
Db 2524 GCAGCCGGCATGCGCCTCAGCCAGGATGGGTGTCACTACGGATGGTGTCCCCGCCAGCAA 2583
Qy 2425 ctgcgcgatccctcttttcctccagaaccctcttccatatggctgccgatggaccaggg 2484
        ||||||||||||||||||||||||
Db 2584 CTGCGCGATCCCTCTTTTCCTTCCAGAACCCTCTTCCATATGGCTGCCGATGGACCAGGG 2643
Qy 2485 ccccagcaggccatggagccagtggcagctgccatgtccacctttccctccatcaccaccac 2544
        ||||||||||||||||||||||||
Db 2644 CCCCAGCAGGCCATGGAGCCAGTGGCAGCTGCCATGTCCACCTTTCCTCCCATCACCACCACCACA 2703
Qy 2545 caccaccatcaccagagggggagaaccccaggggacacttttgcccgttccat 2604
        ||||||||||||||||||||||||
Db 2704 CACCACCATCACCATCACCAGAGGGGAGAACCCCAGGGGACACTTTTGCCCGTTCCAT 2763
Qy 2605 ggctatagcaagaggcaggaccctctctacaggggtccctaagcccctgcggcctccatc 2664
        ||||||||||||||||||||||||
Db 2764 GGCTATAGCAAGAGCAGGAGACCCTCCTTACAGGGGTCCCTAAGCCCCTGCGGCCTCCATC 2823
Qy 2665 cttcacagcacagctacagtgatgagtttggaccctctggtactgattttacccgtcgg 2724
        ||||||||||||||||||||||||
Db 2824 CTTCACAGCACCAGCTACAGTGATGAGTTTGGACCCCTGGGTTACTGATTTTACCCGTCGG 2883
Qy 2725 cagctctcacttcaggacaacctacagcacatgtctcccccgcccagatcaccatcggt 2784
        ||||||||||||||||||||||||
Db 2884 CAGCTCTCACTTCAGGACAACCTACAGCACCTACAGCTCTCCCCGCCCCAGATCACCATCGGT 2943
Qy 2785 cccccagaggccagctccctcaggcccaggggggcagtggtggggcagtggtgggggc 2844
        ||||||||||||||||||||||||
Db 2944 CCCCCAGAGGCCAGCTCCCTCAGGGCCAGGGGGGCAGTGGTGGGGCAGTGGTGGGGGC 3003
Qy 2845 gtgggggccagccagccacctcccctgcagagggccaaatctcagcagttgacagtgagtgct 2904
        ||||||||||||||||||||||||
Db 3004 GTGGGGGCCAGCCAGCCACCTCCCCTGCAGAGGGGCAAATCTCAGCAGTTGACAGTGAGTGCT 3063
Qy 2905 gccagaaaccccggccgtccagcgggaacctattgcagtcccggaaccaagttatggt 2964
        ||||||||||||||||||||||||
Db 3064 GCCCAGAAACCCCGGCCGTCCAGCGGGAACCTATTGCAGTCCCGGAACCAAGTTATGGT 3123
Qy 2965 cctgccgtccacggcaacagagcctcagcaaggggcagcattggggggcagcggggc 3024
        ||||||||||||||||||||||||
Db 3124 CCTGCCGTCCACGGCAACAGAGCCTCAGCAAGAGGGGCAGCATTGGGGGGCAGCGGGGC 3183
Qy 3025 agcggtggcggagggggtgggggctcaagccgccatcaccaagcagcattcccagact 3084
        ||||||||||||||||||||||||
Db 3184 AGCGGTGGCGGCAGGGGGTGGGGGCTCAAGCCGCCTCAAGCCCATTCACCAAGCAGCATTCCCAGACT 3243
Qy 3085 ccattccacggtgaacccacgatgtcgcctgcgtagcctgtagcgggtgtcaat 3144
        ||||||||||||||||||||||||
Db 3244 CCATTCACGGCTGAACCCCACGATGTCCGGCCTCCGGAGCGCAGTGGGGGTGTCCAAT 3303
Qy 3145 atgctcacctgtccgctgacatcgagagtgcacacattgagcggaagagtacaagctg 3204
        ||||||||||||||||||||||||
Db 3304 ATGCTCACCTGTCCGCTGACATCGAGAGTGCACACATTGAGCGGAAGAGTACAAGCTG 3363
Qy 3205 aaggagtactcgaagtccatggacagcatggcctggacaggggtgaaggagtacgaggag 3264
        ||||||||||||||||||||||||
Db 3364 AAGGAGTACTCGAAGTCCATGGACGACAGCAGCGACTGGACAGGGGTGAAGGAGTACGAGGAG 3423
Qy 3265 gagatccactcactgaaggaaaggctacacatgtccaccggaagctgaagagtacgag 3324
```

Right column:

```
        ||||||||||||||||||||
Db 3424 GAGATCCACTCACTGAAGGAAAGGCTACACATGTCCACCGGAAGCTGGAAGAGTACGAG 3483
Qy 3325 cggaggcgtgctgtccccggagagagcagacagaccagcaagccaagccatctgatgcagtcagtacgaccaccccgc 3384
        ||||||||||||||||||||
Db 3484 CGGAGGCGTGCTGTCCCAGGAAGAGCAAGCAGACCAGCAGATCCTGATGCAGTACCAAGCCCGC 3543
Qy 3385 ctggagcagacgagaagcgcttgaggcagcagcaggcagtggagaaggactcccagatcaag 3444
        ||||||||||||||||||||
Db 3544 CTGGAGCAGACAGAGCCAGAAGCGCCTTGAGGCAGCAGCAGGTGGAGAAGGACTC..AGATCAAG 3603
Qy 3445 agcatcattggcaggctgatgtcgtggtggaggagagctcgccggagccaccccgccatg 3504
        ||||||||||||||||||||
Db 3604 AGCATCATTGGCAGGCTGATGTCGTGGTGGAGGAGGAGCTCGCCCGGGAACCACCCGCCATG 3663
Qy 3505 gctgagccgctgcctgaaccaagagagagctgctcgacgctcagagcagaggcagcttcccc 3564
        ||||||||||||||||||||
Db 3664 GCTGAGCCGCTGCCTGAACCCAGAGAGAGGCTCTCGACGCTCGAGAGGAGGCAGCTCCCC 3723
Qy 3565 ccttgggtccaacaaacccgtgtgacgtgccccacctgggaacggctggcccccc 3624
        ||||||||||||||||||||
Db 3724 CCTTGGGTCCAACAAACCCGTGTGACGTGCCGCCCACCTTGGAACGGCCTGGCCCCCC 3783
Qy 3625 cagccccaccccccccccccccccacccgggctcagatcacagaagacggcagtccggaacaccg 3684
        ||||||||||||||||||||
Db 3784 CAGCCCCACCCCCCCCCCCCCCCAGCCTGCAGATCACAGAGAACGGCCAGTTCCGGAACACCG 3843
Qy 3685 cagaccactagcaccagccagcatcacagacctcctgccgtgcaccccgggccca 3744
        ||||||||||||||||||||
Db 3844 CAGACCACTAGCCCACCCAGCATCACAGACCTCTTCCCTGTGCACCCCGGCCCCA 3903
Qy 3745 cccagcgtcacagacctcctccagtgcaccgacctggaacatcaccaccaccagg 3804
        ||||||||||||||||||||
Db 3904 CCCAGCGGTCACAGACCTCTTCCCAGTGCACCGACCCTGGAACATCACCAACCACCAGG 3963
Qy 3805 actggacgtcaccaagggaacagcgggattgtctcccttaacgctcctgggccaccat 3864
        ||||||||||||||||||||
Db 3964 ACTGGACGTCACCAAGGGACCAGCAGCGGCGATTGTCTCCCTTAACGCCTCCTTGGGCACCCAT 4023
Qy 3865 ctgtcaacccccactgctcattccagaggaggaagtggagaccccttccagctgccctctcacc 3924
        ||||||||||||||||||||
Db 4024 CTGTCAACCCCACTGCTCCATCCATCCAGGAGGAGAGTGGGACCCCTCCTCAGCTGCCCTCTCACC 4083
Qy 3925 ccaggacaccacctacccccacacagacccccttcactctgggtgtctatcccccatcct 3981
        ||||||||||||||||||||
Db 4084 CCAGGACACCACTACCCCCACACAGACACCCCCTTCACTCTCGGGGTGCTCATCCCCATCCT 4140
```

```
RESULT 3
AF058789
LOCUS       AF058789    4539 bp    mRNA                ROD       14-SEP-2000
DEFINITION  Rattus norvegicus SynGAP-a mRNA, complete cds.
ACCESSION   AF058789
VERSION     AF058789.2  GI:10122137
KEYWORDS    .
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4539)
AUTHORS     Kim,J.H., Liao,D., Lau,L.F. and Huganir,R.L.
TITLE       SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
            protein family
JOURNAL     Neuron 20 (4), 683-691 (1998)
MEDLINE     98240917
REFERENCE   2  (bases 1 to 4539)
AUTHORS     Kim,J.H. and Huganir,R.L.
TITLE       Direct Submission
JOURNAL     Submitted (09-APR-1998) Neuroscience, Johns Hopkins Medical
            Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
COMMENT     On Sep 14, 2000 this sequence version replaced gi:3065888.
FEATURES             Location/Qualifiers
     source          1..4539
```

```
                /organism="Rattus norvegicus"
                /db_xref="taxon:10116"
                /tissue_type="hippocampus"
CDS             1..3855
                /note="neuronal RasGAP; contains PH domain, C2 domain, GAP
                domain in the N-terminal region, and a T/SXV motif at the
                C-terminus"
                /codon_start=1
                /product="SynGAP-a"
                /protein_id="AAC63510.2"
                /db_xref="GI:10122138"
                /translation="MSYAPPRDVRGPPMHRTQYVHSPYDRPGWNPRFCIISGNQLLML
                DEDEIHPLLIDRRESSRNKLLRRTVSPVEVEGRPHGEHEYHLGRSRRKSVPGGKQYS
                MEAAPAPFRPSQGFLSRRLKSSIKRTKSQPKLDRTSSFRQILPRFRSADHDRARLMQ
                SFKESHSESLLSPSSAAEALENLDEDSIIKPVHSSLGQEFCFEVTTSSGTKCFAC
                RSAAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELPPKKRYYCELCLDDMLY
                ARPTSKPRSASGDTVFWGEHFEFNNLPAVRALRLHLYRDSDKKRKKDKAGYVGLVTVP
                VATLAGRHFTEQWYPVTLPTGSGGSGMGSGGSGGSGGSGKGKGCPAVRLKARYQ
                TMSILPMELYKEFAEYVTNHYRMLCAVLEPALNVKVKGEEVASALVHILQSTGKADPL
                SDMAMSEVDRMEREHLIFRENTLATKAIEEYMRLIGQKYLKDAIGEFIRALYESEEN
                CEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKEVFASWRLRCAERGR
                EDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTLTLIAKVIQNLANFSKF
                TSKEDFLGFMNEFLELEWGSMQQFLYEISNLDTLTNSSSFEGYIDLGRELSTIHALLW
                EVLPQLSKEALLKLGPLPRLLSDISTALRNPNIQRQPSRQSERARSQPMVLRGPSAEM
                QGYMMRDLNSSIDLQSFMARGLNSSMDMARLGHSTVFGEPKPPPPPGGKDLFVSRPPL
                ARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGDPGCGRLNSSSVSNLAAVGDLLH
                SSQASLTAALGLRPAPAGRLSQGSGSSITAAGMRLSQMGVTTDGVPAQQLRIPLSEQN
                PLFHMAADGPGPAGHGGSSGHGPPSSHHHHHHRRGGEPPGDTFAPFHGYSKSED
                LSTGVPKPPAASILHSHYSDEFGPSGTDFTRRQLSLQDNLQHMLSPPQITIGPQRPA
                PSGPGGSGGSGGGGGQPPLQRGKSQOLTVSAAQKPRPSSGNLLQSPEPSYGPAR
                PRQQSLSKEGSIGGSGSGGGGGLKPSITKQHSQTPSTLNPTMPASERTVAWVSNM
                PHLSADIESAHIEREEYKLKEYSKSMDESRLDRVKEYEEIHSLKERLHMSNRKLEEY
                ERRLLSDEEQTSKILMQYQARLEQSEKRLRQQQVEKDSQIKSITGRLMLVEEELRRDH
                PAMAEPLPEPKKRLLDAQLLIR"
BASE COUNT      992 a   1448 c   1290 g    809 t
ORIGIN
Query Match     96.2%;   Score 3828.4;   DB 74;   Length 4539;
Best Local Similarity   97.6%;   Pred. No. 0;
Matches 3940; Conservative    0; Mismatches    16; Indels   82;  Gaps    2;
```

```
Qy      25 gcaccccattcccattgctcccagcaaggcttcctgagccggagcctaaaaagctcc           84
           || |||| | |||     |||   |||||||||||||||||||||||||||||||||
Db     316 GCCCCCGCTGCGCCCTTCCGCCCTTCCGCCAAGGCTTCCTGAGCCGGAGGCTAAAAAGCTCC  375

Qy      85 atcaaacgtacaaagtcacaaccaaacttgaccggaccagcagcttgacagatcctg          144
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     376 ATCAAACGTACAAAGTCACAACCCAAACTTGACCGGACCAGCAGCTTCGACAGATCCTG  435

Qy     145 cctcgcttccgaagtgctgacccgacccgggccccggccggccttcaaggagtct          204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     436 CCTCGCTTCCGAAGTGCTGACCCGACCCGGGCCGCCGGCCGGCCTTCAAGGAGTCT  495

Qy     205 cactcccatgagtccccgctgagtgctgctgaggccctggagctcaacctg          264
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     496 CACTCCCATGAGTCCCCGCTGAGTGCTGCTGAGGCCCTGGAGCTCAACCTG  555

Qy     265 gatgaagactccattatcaagccagtacacagctcacacagtccatcctgggccagagttctgcttt     324
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     556 GATGAAGACTCCATTATCAAGCCAGTACACAGCTCACACAGTCCATCCTGGGCCAGGAGTTCTGCTTT  615

Qy     325 gaggtaacacatcgtctgggacaaaatgtttgcctgtcggtctgcagcgaaagggac          384
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     616 GAGGTAACAACATCGTCTGGGACAAAATGTTTTGCCTGTCGGTCTGCAGCCGCGAAAGGGAC  675

Qy     385 aaatggattgagaatctacagagggctgtgaacccaacaagcaacacagccgccgggta          444
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     676 AAATGGATTGAGAATCTACAGAGGGCTGTGAAACCCAACAAGGACAAAAGCCGCCGCCGGGTA  735

Qy     445 gataacgtgctgaaactatggatcatagaagctcgagagcgtcgagaacgatat          504
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     736 GATAACGTGCTGAAACTATGGATCATAGAAGCTCGAGAGCTCGAGAACGATAT  795
```

```
Qy     505 tactgcgagttatgctgacgacgacatgctctatgtacggacgaccacttccaagcccgctcu          564
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     796 TACTCGCGAGTTATGCTGACGACGACATGCTCTATGTACGGACGACCACTTCCAAGCCCGCTC.  855

Qy     565 gcctcaggagacactgtctttgggcggagcacttgagttaacaacctgctgctgtc          624
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     856 GCCTCAGGAGACACTGTCTTTTGGGCGGAGCACTTGAGTTTAACAACCTGCCTGCTGTC  915

Qy     625 cgggcgctgcgcgctgcatcgtaccgtgactcggacaaaaacgggagaaggacaaggca          684
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     916 CGGGCGCTGCGCGCTGCATCGTGACCGTGACTCGGACAAAAAGCGGAGAAGGACAAGGCA  975

Qy     685 ggctacgttggcctggtgactgttccagtggccacccctggctgggcgccacttcacagag          744
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     976 GGCTACGTTGGCCTGGTGACTGTTCCAGTGGCCACCCCTGGCTGGGCGCCACTTCACAGAG  1035

Qy     745 cagtgtacccgtgacccgtgccaacagaagtggggctctgggggtatgggctcgggg          804
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1036 CAGTGTACCCCGTGACCCGTGCCAACAGGAGTGGGGGCTCTGGGGGTATGGGCTCGGGG  1095

Qy     805 ggagagggggggtcaggggcggctcaggggtcagggggaaagggaggagggtgtcctgctgtg          864
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1096 GGAGAGGGGGGGTCAGGGGCGGCTCAGGGGGCAAAGGGAAAGGAGGAGGCTGTCCTGCTGTG  1155

Qy     865 cggctgaaggccgttaccagacaatgagtatcctgcccatggagctacctataaaggagtt          924
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1156 CGGCTGCAAGGCCCGTTACCAGACAATGAGTATCCTGCCCATGGAGCTATATAAAGGAGTTT  1215

Qy     925 gcagaatatgtgaccaaccactaccgcatgctgtgtgccatgctgctggagcccgcccctcaat          984
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1216 GCAGAATATGTGACCAACCACTACCGCATGCTGTGTGCCATGCTGCTGGAGCCCGCCCTCAAT  1275

Qy     985 gtcaaggggcaagaggaggagtcgctagtgcactgactggttcacatcctgcaacaggcaag          1044
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1276 GTCAAGGGGCAAGAGGAGGAGTCGCTCGTCAGTGCACTGGTTCACATCCTGCAAAGCACAGGCAAG  1335

Qy    1045 gccaaggacttccttcagacatgccatgtcagagtcagggtagaccgttcatggagcgggaa          1104
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1336 GCCAAGGACTTCCTTCAGACACATGCCATGTCAGAGGTAGACCGGTTCATGGAGCGGGAA  1395

Qy    1105 cacctcatattccgcggagaacacgctcgccactaaagccataaagagtatatgagactg          1164
           |||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    1396 CACCTCATATTCCGCGGAGAACACGCTCGCCACTAAAGCCATAGAAGAGTATATGAGACTG  1455

Qy    1165 attggccagaaatacctcaaggatgccattgggagtcatccggctctgtatgaatct          1224
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1456 ATTGGCCAGAAATACCTCAAGGATGCCATTGGGGAGTTCATCCGGGCTCTGTATGAATCT  1515

Qy    1225 gaggagaactgtgaagtagaccccatcaagtgcacagcgtccagtctggcagagcacag          1284
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1516 GAGGAGAACTGTGAAGTAGACCCCATCAAGTGCACAGCGTCCAGTCTGGCAGAGCACCAG  1575

Qy    1285 gccaacctgcgatgtgctgggttggccctgtgcaagtgtgtcaactcccattgcgtg          1344
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1576 GCCAACCTGCGCGATGTGCTGTGCAAGTGTGTCAACTCCCATTGCGTG  1635

Qy    1345 ttcccgagggagctgaagaggtgtttgcatcatggcggctgcgctgcgagagcgggc          1404
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1636 TTCCCGGAGGAGCTGAAGGAGGTGTTTGCATCATGGCGGCTGCGCTGCGAGAGCGGGGC  1695

Qy    1405 cggggagacattgtgacaggctgatcagcgcctcgactccctcctgcggcagggac          1464
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1696 CGGGGAGACATTGTGACAGGCTGATCAGCGCCTCGACTCCCTCCTTCCTGCGGCCGCCG  1755

Qy    1465 gccatcatgtcgcccagtcgtcttggactgatgcaggagtacccagatgagcagacctca          1524
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1756 GCCATCATGTCGCCCCAGTCGTCTTGGACTGATGCAGGAGTACCCAGATGAGCAGACCTCA  1815

Qy    1525 cgaacctcaccctcatcgcaggttatccagaaacctggccaactttccaagtttacc          1584
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1816 CGAACCTCACCCTCATCGCCCAAGGTTATCCAGAAACCTGGCCAACTTTCCAAGTTTACC  1875
```

Fri Jan 19 06:20:01 2001

us-09-294-298-5.rge

```
Qy  1585  tcaaaggagagacttcctgggcttcatgaacgagtttctggactggaatggggttctatg  1644
          |||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  1876  TCAAAGGAGAGGACTTCCTGGGCTTCATGAACGAGTTTCTGGAGCTGGAGTTGGGGTTCTATG  1935

Qy  1645  cagcaattcttgtatgagatatccaacctggacacactgaccaacagcagtttgag  1704
          |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db  1936  CAGCAATTCTTGTATGAGATATCCAACACTGGACACACTGACCAACAGCAGTTTTGAG  1995

Qy  1705  ggctacatagacttgggccggagctctccaacttcacgcctgtctggaggtctg  1764
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  1996  GGCTACATAGACTTGGGCCGGGCCTCTCCAGGGCCAGCTCTCCACACTTCACGCCTGTCTGGGAGGTGCTG  2055

Qy  1765  ccccagctcaagcaaggaagccctctgaagctggcccgctgccccggctcctcagcgac  1824
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2056  CCCCAGCTCAAGCAAGGAAGCCCTCCTGAAGCTGGCCCGGCTGCCCCGGCTCCTCAGCGAC  2115

Qy  1825  atcagcacagacccctgaggaacctaacatccaaagcagcagtccagcagcgagcgc  1884
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2116  ATCAGCACAGACCCCTGAGGAACCCTAACATCCAAAGCAGCAGTCCAGCAGCGAGCGC  2175

Qy  1885  gctcggtctcagcccatggtgctgcgcgggccgtcagccgaaatgcaggggctacatgatg  1944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2176  GCTCGGTCTCAGCCCATGGTGCTGCGCGGGCCGTCAGCCGAGATGCAGGGGCTACATGATG  2235

Qy  1945  cgggaccttcaacagctccatcagcttcagtcctcatggctcgaggcctcaacagctct  2004
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2236  CGGGACCTTCAACAGCTCCATCAGCTTCAGTCCTCATGGCTCGAGGCCTCAACAGCTCT  2295

Qy  2005  atggacatgtctcgctccctcccaaccaagagaaccccgcgcccctccggt  2064
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2296  ATGGACATGTGCTCGCGTCCCTCCCCAACCAAGAGAAACCCCGCGCCCCTCCGGT  2355

Qy  2065  ggggtaaagacctgttctatgtagcggccaccactgccggtctcccccagcatac  2124
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2356  GGGGGTAAAGACCTGTTCTATGTGAGCGGCCCGGTCCTCCCCCAGCATAC  2415

Qy  2125  tgcacgagagctcggacatcacagagccggacagagatgctgagtgtcaacaagagt  2184
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2416  TGCACGAGAGCTCGGACATCACAGAGCCGGACAGAGATGCTGAGTGTCAACAGCAGT  2475

Qy  2185  gtgtcatgtctgctgaccgcagggcgacagggctggggcgcccttaacagcagtagtgtt  2244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2476  GTGTCATGTCTGCTGACCGGGGGCCTGGGGCCGCAGGGGCCTTAACAGCAGTAGTGTT  2535

Qy  2245  tccaacctggcagctgttgggaacctcgttgcactcaagcaggcttcactgacagcagcc  2304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2536  TCCAACCTGGCAGCTGTTGTGGGGACCTGTTGCACTCAAGCAGGCTTCACTGACAGCAGCC  2595

Qy  2305  ttggggttggcggcctgcacctgccggcgcctctccccaaggaggtggctgtcttccatcaca  2364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2596  TTGGGGTTGGCCGCCTGCACCTGCCGCTGCCGGGCGCTCTCCCCAAGGAGGTGGCTCTTCCATCACA  2655

Qy  2365  gcagcgcaggcgccatggacatggggtgtcactacgatggtgtccccgccagcaa  2424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2656  GCAGCGCAGGCCCATGGACATGGGGTGTCACTACGATGGTGTCCCCGCCAGCAA  2715

Qy  2425  ctgcgcatccctctcttcctccagaacctctcttcatatgctgccgatggaccaggg  2484
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2716  CTGCGCATCCCTCTCTTCCTCCAGAACCCTCTTCCATATGCTGCCGATGGACCAGGG  2775

Qy  2485  ccccagcaggccatggagagggggagaacccccaggtgtcacacttcctccatcaccaccac  2544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2776  CCCCAGCAGGCCATGGAGAGGGGAGAACCCCCAGGGGGACACTTCCTCCATCACCACCAC  2835

Qy  2545  caccaccatcaccatccaccatccacaggggaacactttgcccgttccat  2604
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2836  CACCACCATCACCATCACCACCATCACCGGGGACACTTTTGCCCGTTCCAT  2895

Qy  2605  ggctatagcaagacgcgaggacctctctacaggggtccctaagcccctgcggcctccatc  2664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2896  GGCTATAGCAAGACGCGAGGACCCTCTCTACAGGGGTCCCTAAGCCCCTGCGGCCTCCATC  2955

Qy  2665  cttcacagccacagctacagtgatgagtttggaccctctggtactgatttacccgtcgg  2724
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2956  CTTCACAGCCACAGCTACAGTGATGAGTTTGGACCCTCTGGTACTGATTTTACCCGTCGG  3015

Qy  2725  cagctctcacttcaggacaacctacagcacatgtctccccgcccagatcaccatcggt  2784
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3016  CAGCTCTCACTTCACTGCAGACAACACTACAGCACATGTCTCCGCGCCCAGATCACCATCGGT  3075

Qy  2785  ccccgagaggcagagctccctcaggccaggagggggcagtggtggggcagtggtgggggc  2844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3076  CCCCAGAGGGCCAGCTCCCTCAGGGCCAGGAGGGGGCAGTGGTGGGGGCAGTGGTGGGGGC  3135

Qy  2845  ggtgggggcagccagccacctccttgcagagggcaaatctcagcagttgacagtgagtgct  2904
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3136  GGTGGGGGCCAGCCACCTCCCTTGCAGAGGGGGCAAATCTCAGCAGTTGACAGTGAGTGCT  3195

Qy  2905  gcccagaaaccccgtccagcgcggaaccctattgcgtccccggaaccaagttatggt  2964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3196  GCCCAGAAACCCCGGCGTCCAGCGCGGAACCTATTGCGAGTCCCCGGAACCAAGTTATGGT  3255

Qy  2965  cctgcccgtccacgcaacagagctcagcaaagaggcagcattggggcagcgggggc  3024
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3256  CCTGCCCGTCCACGGCCAACGGGCTCAGCAACAGAGCTCAGCAAAGAGGGGCAGCGGGGGC  3315

Qy  3025  agcggtggcggagggggtgggggctcaagccctccatcaccaagcagcagcattcccagact  3084
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3316  AGCGGTGGCGGAGGGGGTGGGGGCTCAAGCCCTCCATCACCAAGCAGCAGCATTCCCAGACT  3375

Qy  3085  ccatccacgctgaaccccacgatgccgggcctccggagcggactgtagcctggtgtccaat  3144
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3376  CCATCCACGCTGACCCCACGATGCCGGCCTCCGGAGCGGCCGGACTGTAGCTCTGGTGTCCAAT  3435

Qy  3145  atgcctcacctgtccgctgacatcgagagtgcacacattgacgggaagagtacaagctg  3204
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3436  ATGCCTCACCTGTCCGCTGACATCGAGAGTGCACACATTGAGCGCGAGAGTACAAGCTG  3495

Qy  3205  aaggagtactcgaagtccatggacagagatccgagtgagtgtgagcagggtgaaggagtacgagag  3264
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3496  AAGGAGTACTCGAAGTCCATGGACAGAGATCCGACGAGCAGGGTGAAGGAGTACGAGGAG  3555

Qy  3265  gagatccactcactgaaggaaagctacacacattccaaccgaaggctggaaggtacgag  3324
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3556  GAGATCCACTCACTGAAGGGAAAGGTACACATGTCCAACCGGAGAGCTGGAAGAGTACGAG  3615

Qy  3325  cggaggctgctgtcgtccaggaagagagaccacagaccaggacttcatcagtcagaagcagcagc  3384
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3616  CGGAGGCTGCTGTCGTCCAGGAAGAGAGCCAGCAGACCAGGACTTCATGCAGTCGAACCAAGGAGCAGCCCGC  3675

Qy  3385  ctggagcagagcgagaagcttgagcgacagcaggcttgagggctggagaaggactcccagatcaag  3444
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3676  CTGGAGCAGAGCGAGAAGCGCTTGAGCGACAGCAGGCTGGAGAAGGACTCCCAGATCAAG  3735

Qy  3445  agcatcattggcaggctgatgctggttgaggaggagagctgcgcgggaccacccgccatg  3504
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3736  AGCATCATTGGCAGGCTGATGCTGTGGTGGAGGAGGAGTGCGCCGGGACCACCCGCCATG  3795

Qy  3505  gctgagcgcgtgcctgaaccaagaaagctctgctcgacgctca ------------  3548
          |||||||||||||||||||||||||||||||||||||||||||
Db  3796  GCTGAGCGCCGCTGCCTGAACCCAAGAAAGAGGCTGCTCGCAGCTCTCATCAGTAA  3855

Qy  3549  ------  3548
          ------
Db  3856  TTCTCCTGGTTCCGGCTTTGACCACGGGCCGGAGGACACTTGGGGGAGGTGACTCCGGACCAC  3915

Qy  3549  ----gagaggcagcttcccccttggtcacaacaaacccggttgtacggtggcccac  3603
          ||||||||||||||||||||||||||||||||||||||||||||
Db  3916  TGCAGGAGGAGCGGCAGCTTCGCCAACAAACCCGCTTGGGGTCCAACAAACCCGCTGTACGCGTGGCCCCCAC  3975

Qy  3604  cttggaacggcctggcccccagccccaccccccaccccggctgcagatcacagaga  3663
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3976  CTTGGAACGGCCTGGCCCGCCCCCCCCCGCCCCGCTGCCGGGCTGCAGAATCACAGAGA  4035

Qy  3664  acggcagttccggaacacgcgagaccactagccaccagcatcacagacctccttccc  3723
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   4036  ACGGCGAGTTCCGGAACACCGCAGACCACTAGCCCACCCAGCATCACAGACCTCCTTCCC  4095
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3724  tgtgcaccctaccccgccaccagcgtcacagacctcttccagtgcaccggaccct     3783
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4096  TGTGCACCCTACCCCGCCACCCAGCGTCACAGACCTCCTTCCCAGTGCACCCGACCCT   4155
Qy   3784  ggaacatcaccaaccaccaggactgacgtcaccaagggacagcgggattgtctcccttа  3843
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4156  GGAACATCACCAACCACCAGGACGTGACGTCACCAAGGGACAGCGGGATTGTCTCCCTTA  4215
Qy   3844  acgcctccttgggcaccatctgtcaaccccactgctccattccaggagggagagtggg    3903
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4216  ACGCCTCCTTGGGCACGC-TCTGTCAACCCCACTGCTCCATTCCAGGAGGAGAGTGGG    4274
Qy   3904  acctcagtcgcctctcacccagaacacacctcaccccacacagacccctccactctg    3963
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4275  ACCCTCAGCGCGTGCCCTCACCCCAGGACACACCTCACCCCACACAGACCCCTTCACTCTG  4334
Qy   3964  ggtgctatcccatct      3981
            |||||||||||||||
Db   4335  GGGTGCTATCCCCATCCT     4352
```

RESULT 4
AF048976
LOCUS   AF048976   4063 bp   mRNA   ROD   27-MAR-1998
DEFINITION Rattus norvegicus synaptic ras GTPase-activating protein p135
            SynGAP mRNA, complete cds.
ACCESSION   AF048976
VERSION   AF048976.1 GI:2935447
KEYWORDS   .
SOURCE   Norway rat.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 4063)
  AUTHORS   Chen,H.-J. and Kennedy,M.B.
  TITLE   Identification and cloning of a novel 130 kd protein containing a
            ras GTPase-activating domain from the rat forebrain postsynaptic
            density
  JOURNAL   (in) SOC. NEUROSCI. ABSTR.: 1466:(1997)
REFERENCE   2 (bases 1 to 4063)
  AUTHORS   Chen,H.-J., Rojas-Soto,M. and Kennedy,M.B.
  TITLE   A synaptic Ras GTPase-activating protein (p135 SynGAP) inhibited by
            CaM kinase II
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 4063)
  AUTHORS   Chen,H.-J. and Kennedy,M.B.
  TITLE   Direct Submission
  JOURNAL   Submitted (17-FEB-1998) Division of Biology, California Institute
            of Technology, 1200 E. California Blvd., MC 216-76, Pasadena, CA
            91125, USA
FEATURES   Location/Qualifiers
     source   1..4063
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     CDS   122..4003
                     /function="prominent substrate for endogenous CaMKII; can
                     activate intrinsic ras GTPase activity"
                     /note="synaptic ras-GAP; N-terminal encodes putative PH
                     domain, C2 domain, and ras-GAP domain; C-terminal encodes
                     proline-rich region, stretch of 10 histidine residues, and
                     t/SXV motif; enriched in the forebrain postsynaptic
                     density fraction"
                     /codon_start=1
                     /product="synaptic ras GTPase-activating protein p135
                     SynGAP"
                     /protein_id="AAC08071.1"
                     /db_xref="GI:2935448"
                     /translation="MSYAPFRDVRGPPMHRTQYVHSPYDRPGWNPRPCIISGNQLIML
                     DEDEITHPLLIRDRRSESSRNKLLRRTVSVPVEGRPHGEHEYHLGRSRRKSVPGGKYS
```

```
MEAAPAAPFRPSQGFLSRRLKSSIRRTKSQPKLDRTSSFRQILPKFRSADHDRARLMQ  4095
SFKESHSHESLLSPSSAAEAELNLDEDSIIKPVHSSILGQEFCFEVTTSSGTKCFAC  3783
RSAAERDKWIENLQRAVKPNKDNSRRVDNVLKIWIIEARELPPHKRYYCELCLDDMLY   4155
ARTTSKPRSASGDTVPWGEHFEFNNLPAVRALRLHLYRDSDKKKKDKAGYVGLVTVP
VATLAGRHFTEQWYPVTLPTGSGSGSGMGSGGGSGGGSGGGSGGGGKGSGCPAVRLKARYQ  84
TMSLPMELYKEFAEYVTNHYRMLCAVLEPALNVKGEVASALVHILQSTGKAKDFL
SDMAMSEVDRFMEREHLJFRENTLAPTKAIEEYMRLIGQKYLKDAIGEFIRALVESEEN  3843
CEVPDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKEVFASWRLRCAERGR
EDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDQTSRTLTLIAKVIQNLANFSKF  4215
TSKEDFLGFMNEFLELEWGSMQQFLYEISNLDTLTNSSFPEGYIDLGRELSTLHALLW
EVILPQLSKEALLKLGPLPRLLSDISTALRNPNIQRPSRQSERARSQPMVLRGPSAEM  3903
QGYMMRDLNSSIDLQSFMARGLNSSMDMARLPSPTKEKPPPPPGGGKDLFYVSRPPL
ARSSPAVCTSSSDITEPEQKMLSVNKSVSMLDLQGDPGGRLNSSSVSNLAAVGDLLH  4274
SSQASLTRALGLRPAPAGRLSQGSGSSSITRAGMRLSQMGVTTDGVPAQOLRIPLSFQN
PLFHMAADGPGPPAGHGGSSGHGPPSSHHHHHHHHHHRGGEPPGDTFAPFHGYSKSED  3963
LSTGVPKPPAASILHSHSYSDEFGPSGTDFTRROLSLQDNLQHMLSPPOITIGPQRPA
PSGPGGSGGGSGGSGGGGQPPLQRGKSQOLTVSAAQKPRPSSGNLLQSPEPSYGPAR  4334
PRQQSLSKEGSTGSSGGSGGGGGGLKPSITKQHSOTPSTLNPTMPASERTVAWVSNM
PHLSADIESAHIEREEYKLKEYSKSMDESRLDRVKEYEEEIHSLKERLHMSNRKLEEY  3981
ERRLLSQEQTSKILMQYQARLEQSEKRLRQQQVERDSQIKSIIGRLMLVEEELRRDH
PAMAEPLPEPKKRLLLDAQRGSFPPWVQQTRV"     4352
```

```
BASE COUNT     880 a      1273 c      1177 g      733 t
ORIGIN
```

Query Match   90.5%; Score 3601.4; DB 74; Length 4063;
Best Local Similarity   99.6%; Pred. NO. 0;
Matches 3611; Conservative   0; Mismatches 16; Indels   0; Gaps   0;

```
Qy    25  gcacccatttccattgtctcccagcaaggcttcctgagccggagctaaaagctcc      84
           || ||  || |  ||||   ||  ||||||||||||||||||||||||||||||||
Db   437  GCCCCCGCTGCGCCCTTCCGCCCTGCCAAGGCTTCCTGAGCCGGAGCTAAAAGCTCC   496
Qy    85  atcaaacgtacaaagtcacaacccaaactgaccggaccagcagcttcgacagatcctg  144
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   497  ATCAAACGTACAAAGTCACAACCCAAACTGACCGGACCAGCAGCTTCGACAGATCCTG   556
Qy   145  cctcgcttccgaagtgctgaccatgacgacggaccggggccggcgtgatgcaaggagtct  204
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   557  CCTCGCTTCCGAAGTGCTGACCATGACCGGCCGGCCCGGCTGATGCAGAGCTTCAAGGAGTCT  616
Qy   205  cactccatgagtccctgctgagtcccagcagtgctgctgaggccctggagctcaacctg   264
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   617  CACTCCCATGAGTCCCTGCTGAGTCCCAGCAGTGCTGCTGAGGCCCTGGAGCTCAACCTG  676
Qy   265  gatgaagactccattatcaagccagtacacagctcacagctacacagccatctggccaggagttctgcttt  324
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   677  GATGAAGACTCCATTATCAAGCCAGTACACAGCTCACAGCTACACCTCCGGCCAGGAGTTCTGCTTT  736
Qy   325  gaggtaacaacacctgtctggacaaaatgtttttgcctgtcggtctgagccgaaagggac  384
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   737  GAGGTAACAACATCGTCTGGACAAAAATGTTTTGCCTGTCGGTCTGCAGCCGAAAGGGAC  796
Qy   385  aaatgattgagaatctacagagggctgtgaaccaacaagacaacagcgccgggta     444
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   797  AAATGATTGAGAATCTACAGAGGGCTGTGAAACCAACAACAAGGACAACAGCGCCGCCGGGTA   856
Qy   445  gataacgtgctgaaactatggatcatagaagtcgagagctgccccccagaagcgatat     504
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   857  GATAACGTGCTGAAACTATGGATCATAGAAGTCTCGAGAGCTGCCCCCCCAGAAGCCATAT   916
Qy   505  tactgcgagttatgcctggacgacatgctctatgcacggaccacttccaagcccgctca   564
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   917  TACTGCGAGTTATGCCTGGACGACATGCTCTATGCACGGACCACTTCCAAGCCCCCTCA   976
Qy   565  gcctcaggagacactgtctttggggcgagcacttcgagtttaacaacctgctgctgtc   624
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   977  GCCTCAGGAGACACTGTCTTTGGGGGCGAGCACTTCGAGTTTAACAACCTGCTGCTGTC  1036
Qy   625  cgggcgctgcggctgcatcgtaccgtgactggacaaaaagcggaagaagggacaaggca   684
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1037  CGGGCCTGCGCGCTGCATCGTGACTGGACAAAAAGCGGAGAAGGACAAGGCA  1096
```

```
Qy    685  ggctacgttggcctggtgactgttccagtggccaccctggctggggccactcacagag  744
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1097  GGCTACGTTGGCCTGGTGACTGTTCCAGTGGCCACCCTGGCTGGGGCCACTTCACAGAG 1156

Qy    745  cagtggtaccccgtgacctgccaacaggaagtggggctctggggtatgggctcgggg    804
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1157  CAGTGGTACCCCGTGACCCTGCCAACAGGAACCCTGGGGTATGGGCTCGGGG       1216

Qy    805  ggaggaggggggtcaggggcggctcaggggcaaagggaagagggctgtgtcctgctgtg   864
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   1217  GGAGGAGGGGGTCAGGGGCGGCGGCTCAGGGGCAAAGGGAAAGGAGGCTGTCCTGCTGTG 1276

Qy    865  cggctgaaggcccgttaccagacaatgagtatcctgcccatggagctatataagggagttt   924
           ||||||||||||||||||||||||||||||||||||||
Db   1277  CGGCTGAAGGCCCGTTAACCAGACAATGAGTATCCTGCCCATGGAGCTATATAAGGAGTT  1336

Qy    925  gcagaatatgtgaccaaccactaccgcatgctgttgtgccgtgctggagcccgccctcaat   984
           |||||||||||||||||||||||||||||||||||||||
Db   1337  GCAGAATATGTGACCAACCACTACCGCATGCTGCTCGCGTCTGGAGCCCGCCCTCAAT   1396

Qy    985  gtcaaggcaaggaggagtcgtagtgcactgttcacatcctgcaaagcacagcaag      1044
           |||||||||||||||||||||||||||||||||||||||
Db   1397  GTCAGGGCAAGGAGGAGTCGTAGTGCACTGTTCACATCCTGCAAAGCACAGGCAAG   1456

Qy   1045  gccaaggacttcctttcgacatggccatgtcagaggtagaccggttcatggagcgggaa  1104
           |||||||||||||||||||||||||||||||||||||||
Db   1457  GCCAAGGACTTCCTTTCAGACATGGCCATGTCAGAGGTAGACAGAGGTTCATGGAGCGGGAA 1516

Qy   1105  cacctcatattccgcgagaacacgctcgccactaaagccatagaagagtatatgagactg  1164
           |||||||||||||||||||||||||||||||||||||||
Db   1517  CACCTCATATTCCGCGAGAACACGCTCGCCACTAAAGCCATAGAAGAGTATATGAGACTG 1576

Qy   1165  attggccagaaatacctcaaggatgccattgggagttcatccgggctctgtatgaatct  1224
           |||||||||||||||||||||||||||||||||||||||
Db   1577  ATTGGCCAGAAATACCTCAAGGATGCCATTGGGAGTTCATCCGGGCTCTGTATGAATCT 1636

Qy   1225  gaggagaactgtgaagtagagccccatcaagtcgacacgctcagtgctggagagcaccag  1284
           |||||||||||||||||||||||||||||||||||||||
Db   1637  GAGGAGAACTGTGAAGTAGAGCCCCATCAAGTCGACACGCTCAGTGCTGGAGAGCACCAG 1696

Qy   1285  gccaacctgcggatgtgtcgtgtgagttggccctgcaaggtggtcaactccattgcgtg  1344
           |||||||||||||||||||||||||||||||||||||||
Db   1697  GCCAACCTGCGGATGTGTCGTGTGAGTTGGCCCCTGTGTCAAGGTGGTCAACTCCATTGCGTG 1756

Qy   1345  ttcccgagggagctgaaggaggtgtttgcatcatggcgcgctgcgctgtgcagagcggggc  1404
           |||||||||||||||||||||||||||||||||||||||
Db   1757  TTCCCGAGGGAGCTGAAGGAGGTGTTTGCATCATGGCGCGCTGCGCTGTGCAGAGCGGGGC 1816

Qy   1405  cggggagacattgctgacagctgtatcagcggcctcgctcttcctgcgcttcctctgcccg  1464
           |||||||||||||||||||||||||||||||||||||||
Db   1817  CGGGGAGACATTGCTGACAGCTGTATCAGCGGCCTCGCTCTTCCTGCGCTTCCTCTGCCCG 1876

Qy   1465  gccatcatgtcgcccagctgtttggactgagtagtcaaggtggtcaactgacagacctca  1524
           |||||||||||||||||||||||||||||||||||||||
Db   1877  GCCATCATGTCGCCCAGTCTGTTGGACTGAGTTGGACTGATCAGGAGTACCAGATGAGCAGGACCTCA 1936

Qy   1525  cgaacctcacccatgccaaggttatccaaggctggccaacttttcaagtttacc      1584
           |||||||||||||||||||||||||||||||||||||
Db   1937  TTCCCGGAGGGAGCTGAAGGAGGTGTTTGCATCATGGCGCGCTCATGCGGCTCCAAGTTTCCAAGTTTACC 1996

Qy   1585  tcaaaggaggactcctggcttcatgacgagtttctggaatggggttctatg     1644
           ||||||||||||||||||||||||||||||||||||
Db   1997  TCAAAGGAGGACTTCCTGGCTTCATGAACGAGTTTCTGGAGTGGGGTTCTATG    2056

Qy   1645  cagcaattccttgtgatgagatatccaacctgaccacagcaacagcacagcagtttttgag  1704
           |||||||||||||||||||||||||||||||||||||||
Db   2057  CAGCAATTCTTTGTGATGAGATATCCAACCTTCAGGACACAACAGCACACAGCAGTTTTTGAG 2116

Qy   1705  ggctacatagactctgggccgagctcacgccctgctctcaccgccctgctggagagtgctg  1764
           |||||||||||||||||||||||||||||||||||||||
Db   2117  GGCTTACATAGACTTGGGCCGGGCCTCCACTTGGGCCCGAGCTCCTGCTCTCGGGAGGTGCTG 2176

Qy   1765  cccccagctcagcaaggaagccctcctgaagctggccccgctgccccgcctgcctcagcgac  1824
           |||||||||||||||||||||||||||||||||||||||
```

```
Db   2177  CCCAGCTCAGCAAGGAAGCCCTTCCTGAAGCTGGCCCCGCTGCCCGGGCTCCTCAGCGAC 2236

Qy   1825  atcagcacagcccctgagaacctaacatccaaaggaggagcagccggagcccgccagagcgagcgc 1884
           |||||||||||||||||||||||||||||||||||||||
Db   2237  ATCAGCACAGCCCCTGAGGACCCCTAACATCCAAACATCCAAAGGCAGCCGCCGCCAGCGCGGGAGCGC 2296

Qy   1885  gctcggtctcagccatggccatgtgctgcgggggcgtcacgcagcgaaatgcagggctcaacatgatg 1944
           |||||||||||||||||||||||||||||||||||||||
Db   2297  GCTCGGTCTCAGCCATGGCCATGTGCTGCGGGGCCCTCAGCCGACATGCAGATGTCGAGGGCTACAATGATG 2356

Qy   1945  cgggacctcaacactccatcgacttcagtccttcagtcctcgaggcctcaacagctct 2004
           |||||||||||||||||||||||||||||||||||||||
Db   2357  GGGACCTCAACACTCCATCGACTTCCATTCAGTCCTTCATCAGTGGCTCGAGGCCTCAACAGCTCT 2416

Qy   2005  atggacatggctcgcctccctccccaccaaggagaaacccccgccgcccctccccgggt 2064
           |||||||||||||||||||||||||||||||||||||||
Db   2417  ATGGACATGGCTCGCCTGCCTCCCTCCCCAACCAAGGAGAAAACCCCCGGCCGCCCTCCCCGGT 2476

Qy   2065  gggggtaaagaccgttctcatgtgagccggccacactggccggtcctcccagcatac 2124
           |||||||||||||||||||||||||||||||||||||||
Db   2477  GGGGTAAAGACCGTTCTCTATGTGAGCCGGCCGGCCCACTGGCCGGTCTCCCCAGCATAC 2536

Qy   2125  tgcacgacagctcggacatcacagagaccggagcagaagatgctggatgctgaatgcagcgcc 2184
           |||||||||||||||||||||||||||||||||||||||
Db   2537  TGCACGGACAGCTCGGACATCACAGAGACCGGAGCAGAAGATGCTGACGTGTCAACAAGAGT 2596

Qy   2185  gtgtccatgtgacctgacctgcagggcgcagggcctggggcctgggggcctggtgatgtt 2244
           |||||||||||||||||||||||||||||||||||||||
Db   2597  GTGTCCATGTGACCTGACCTGCAGGCGCAGGGGCCTGGGGGCCTGGGGGCCTTAACAGCAGTAGTGTT 2656

Qy   2245  tccaacctgcagctgtgtgggacactgttgactcaagccaggcttcactgacagcagcc 2304
           |||||||||||||||||||||||||||||||||||||||
Db   2657  TCCAACCTGGCAGCTGTGTTGGGACCTGTTGCACTCAAGCCAGGCTTCACTGACAGCAGCC 2716

Qy   2305  ttggggtttgcgggctgcacctgccggcggcgctctctccccaaggagtggctcttccatcaca 2364
           |||||||||||||||||||||||||||||||||||||||
Db   2717  TTGGGGTTTGCGGGCTGCACCTGCCGGGCGGCGGCCCTGCTCCCCAAGGGAGTGGCTCTTCCATCACA 2776

Qy   2365  gcagccggcgatggcgatggcctcagcagatggggtgtcactacggatgatgtgtgtccagcaa 2424
           |||||||||||||||||||||||||||||||||||||||
Db   2777  GCAGCCGGCGATGGCGATGGCCTCAGCCAGATGGTGTCACTACGGATGGTGTCCCCGCCAGCAA 2836

Qy   2425  ctgcgcatccctctttcctccagaacctctccttccatatggctgccgatggaccaggg 2484
           |||||||||||||||||||||||||||||||||||||||
Db   2837  CTGCGCCATCCCTCTCTTTCCTTCCTTCCAGAACCCTCTCCATATGGCTGCCGATGGACCAGGG 2896

Qy   2485  cccccagcaggcagcatgaggagcagtacagtggccatggtcacccttcctccatcaccaccac 2544
           |||||||||||||||||||||||||||||||||||||||
Db   2897  CCCCAGCAGGCCATGAGGAGGCAGTAGCAGGCCATGGCCACCTTCCACCTTCCCATCACCACCAC 2956

Qy   2545  caccaccatcaacctcaccagagggagaacccccagggacactttgcccgttccat 2604
           |||||||||||||||||||||||||||||||||||||||
Db   2957  CACCACCATCACCATCAACCTCACCAGAGGGAGAACCCCCAGGGGACAACACTTTTGCCCCGTTCCAT 3016

Qy   2605  ggctatacaagagcgaggacctctcacagggtcctaagcccctgcggccccatc 2664
           |||||||||||||||||||||||||||||||||||||||
Db   3017  GGCTATAGCAAGAGCGAGGACCTCTCTACAGGGGTCCTAAGCCCCTGCGGCCCCATC 3076

Qy   2665  cttcacagccacagctacagtgatgatgagtttgaaccttctgtgtactgatttttaccgtcgg 2724
           |||||||||||||||||||||||||||||||||||||||
Db   3077  CTTCACAGCCACAGCCACAGTGATGATGAGTTTGGAACCTCTGTGGTACTGATTTTTACCCGTCGG 3136

Qy   2725  cagctctcacttcaggacaacctacagacacatgtctctccccccagatcaccaccggt 2784
           |||||||||||||||||||||||||||||||||||||||
Db   3137  CAGCTCTCACTTCAGGGACAACCTACAGCACACATGTCTCTCCCCGCCCCCAGATCACCATCGGT 3196

Qy   2785  ccccagaggccagctcccccaggcaggaggggcagtgtggggcagtggtggggggc 2844
           |||||||||||||||||||||||||||||||||||||||
Db   3197  CCCCAGAGGCCAGCTCCCCAGGGCAGGAGGGGCAGTGGTGGGGCAGTGGTGGGGGGC 3256

Qy   2845  ggtggggggccagcagccaccctccccttgacagagggcaaatccaggttgacagtgagtgact 2904
           |||||||||||||||||||||||||||||||||||||||
```

```
Db   3257  GGTGGGGGCCAGCCACCTCCCTTGCAGAGGGCAAATCTCAGCAGTTGACAGTGAGTGCT  3316
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2905  gcccagaaaccccggccgtccagcggaacctattgcagtccccggaaccaagttatggt  2964
Db   3317  GCCCAGAAACCCCGGCCGTCCAGCGGAACCTATTGCAGTCCCCGGAACCAAGTTATGGT  3376
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   2965  cctgccccgtccacggcaacagagacctcagcagaagggcagcattggggcagcggggc  3024
Db   3377  CCTGCCCCGTCCACGGCAACAGAGACCTCAGCAGAAGGGCAGCATTGGGGGCAGCGGGGGC  3436
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3025  agcggtggcggaggggtggggggctcaagcctcatcaccaagcagcattccagact  3084
Db   3437  AGCGGTGGCGGAGGGGGGTGGGGGGCTCAAGCCCTCCATCACCAAGCAGCATTCCCAGACT  3496
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3085  ccatccacgctgaccccacgatgccggcctcggacggactgtagcctggtgtccaat  3144
Db   3497  CCATCCACGCTGAACCCCACGATGCCGGCCTCGGACGGACTGTAGCCTGGTGTCCAAT  3556
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3145  atgctcacctgtccgctgacatcggagtgcacacattgagcgggaagagtacaagctg  3204
Db   3557  ATGCTCACCTGTCCGCTGACATCGGAGTGCACACATTGAGCGGGAAGAGTACAAGCTG  3616
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3205  aaggagtactcgagtccatggacgagagcgactgacaggggtgaaggagtacgaggag  3264
Db   3617  AAGGAGTACTCGAGTCCATGGACGAGAGCGACTGACAGGGGTCAAGGAGTACGAGGAG  3676
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3265  gagatccactcactgaaggaaagctacacatgtccaaccggaagctggaagagtacag  3324
Db   3677  GAGATCCACTCACTGAAGGAAAGCTACACATGTCCAACCGGAAGCTGGAAGAGTACGAG  3736
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3325  cggaggctgctgtccaggagagcagacaagatcctgatgcagtcagtaccaagcccgc  3384
Db   3737  CGGAGGCTGCTGTCCCAGGAAGAGCAGACAAGATCCTGATGCAGTCAGTACCAAGCCCGC  3796
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3385  ctggagcagagcagaagcgcttgagcagcagcaggtgagaaggagtcccagatcaag  3444
Db   3797  CTGGAGCAGAGCGAGAAGCGCTTGAGCGAGCAGCAGGTGGAGAAGGAGCTCCCAGATCAAG  3856
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3445  agcatcattggcaggctgatgctggtggagagggactgcgccggaccaccccgcatg  3504
Db   3857  AGCATCATTGGCAGGCTGGTGGTGGGAGGAGGCTGCCCCGGGACCACCCCGCCATG  3916
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3505  gctcagagccgctgctgaacccaagagagaggctgctcgacgctcagagagcagcttcccc  3564
Db   3917  GCTCAGAGCCGCTGCTGAACCCAAGAGAGAGGCTGCTCGACGCTCAGAGGCAGCTTCCCC  3976
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3565  cctggggtccaacaaaccgcgtgtgacggtggcccaccttggaacggcctggcccccc  3624
Db   3977  CCTTGGGTCCAACAAACCCCGTGTGACGGTGGCCCACCTTGGAACGGCCTGGCCCCCCC  4036
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3625  cagcccccacccccccaccccggctgc  3651
Db   4037  CAGCCCCACCCCCCCACCCCGGGTGC  4063
```

Query Match   89.9%; Score 3579.6; DB 74; Length 4801;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 3638; Conservative 0; Mismatches 14; Indels 20; Gaps 3;

```
Qy     28  cccccattccatttgctccccagcaaggcttcctgagccggaggctaaaaagctccatc  87
            ||||||||||| |||  ||| ||| ||||||||||||||||||||||||||||||||||
Db   1136  CCCGCTGCCGCCCCTTCCGGCCCTCGCAAGGCTTCCTGAGCCGGAGGCTAAAAAGCTCCATC  1195
Qy     88  aaacgtacaaagtcacaaccaaacttgaccggaccagcagcttggacagatcctgcct  147
            ||||||||||| |||||||||||||||  ||||||||||||||||||||||||||||||
Db   1196  AAACGTACAAAGTCACAAACCAACCCAAACTTGACCGGACCAGCAGCTTCGACAGATCCTGCCT  1255
Qy    148  cgcttccagagtgctgaccatgaccggcccggtcagtgcagagcttcaggagtctcac  207
            ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   1256  CGCTTCCGAGTGCTGACCATGACCGGCCCCGGCTAAGGTTCAGGAGCTTCAGGAGTCTCAC  1315
Qy    208  tcccatgacaaagtcacaaccaaacttgaccggacagagtgctgtgagcgtgagctcaacctggat  267
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   1316  TCCCATGACAAAGTCCACAACCAAACTTGACCGGACCAGCAGCTTGAGCTGCTGTGAGCTCAACCTGGAT  1375
Qy    268  gaagactcattatcaagccagtacacagctccatcctggccaggagttctgcttgag  327
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1376  GAAGACTCCATTATCAAGCCAGTACACAGCTCCATCCTCGGCCAGGAGTTCTGCTTTGAG  1435
Qy    328  gtaacaacatcgtctggaacaaaatgtttgcctgctcggtctgcagccgaaagggacaaa  387
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1436 GTAACAACATCGTCTGGGACAAAATGTTTTGCCTGTCGGTCTCGAGCGCCGAAAGGGACAAA 1495

Qy   388 tggattgagaattcacagagggctgtgaaacccacacaaggacaacagcgcgcgggtagat  447
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1496 TGGATTGAGAATTCACAGAGGGCTGTGAAACCCACACAAGGACAACAGCCACCGGGTAGAT 1555

Qy   448 aacgtgctgaaactatggatcatagaagctcgagagctcgcccccaagaagcgatattac  507
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1556 AACCTGCTGAAACTATGGATCATAGAAGCTCGAGAGCTCGCCCCCAAGAAGCGATATTAC 1615

Qy   508 tgcgagttatgctggacgacatgtctctatgcacggaccacttccaagccccgctcagcc  567
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1616 TGCGAGTTATGCTGGACGACATGTCTATGCACGGACCACTTCCAAGCCCCGCTCAGCC 1675

Qy   568 tcaggagacactgtctttgggcgagcacttcaacaacctgctgctgtccgg  627
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1676 TCAGGAGACACTGTCTTTGGGCGAGCACTTCAACAACCTGCCTGCTGTCCGG 1735

Qy   628 gcgctgcggctgcatctgtaccgtgactcggacaaaaacggaagagaaggacaaggcaggc  687
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1736 GCGCTGCGGCTGCATCGTGTACCGTGACTCGGACAAAAAGCGGAAGGACAAGGCAGGC 1795

Qy   688 tacgttggcctgtgactgttccagtggccacccctggccggcctggcctggcctggcc  747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1796 TACGTTGGCCTGTGTGACTGTTCCAGTGGCCACCCTGGCCTGGCCTGGCCTGGCC 1855

Qy   748 tggtacccgtgaccctgccaacaggaagtggggctctgggggtatgggctcgggggga  807
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1856 TGGTACCCGTGACCCTGCCAACAGAGAGTGGGGCTCTGGGGTATGGGCTCGGGGGGA 1915

Qy   808 ggaggggggtcaggggcggctcaggggcaaaggaaggagctgtctgctgtgtcagg  867
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1916 GGAGGGGGTCAGGGGCGGCTCAGGGGCAAAGGGGCAAAGGGAGGCTGTCTGCTGTGCGGG 1975

Qy   868 ctgaaggcccgttacagacaatgagtatcctgccatggagctatataagggagttgca  927
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1976 CTGAAGGCCCGTTACCAGACAATGAGTATCCTGCCCATGGAGCTATATAAGGAGTTTGCA 2035

Qy   928 gaatatgtgaccaaccactaccgcatgtctgtgtcgcgtgtgctgagccccctcaatgtc  987
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2036 GAATATGTGACCAACCACTACCGCATGTCTGTGTGCCGTGCTGTGCTGAGCCCCCTCAATGTC 2095

Qy   988 aagggcaaggagaggagtcgctagtcgactgttcacatctgttcaaaagcacaggcaaggcc 1047
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2096 AAGGGCAAGGAGAGGAGGTCGCTAGTGCACTGTTCACATCCTGCAAAGCACAGGCAAGGCC 2155

Qy  1048 aaggacttccttcagacatgccatgtgaggtagaccggttcatggagcgggaacac 1107
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2156 AAGGACTTCCTTCAGACATGCCGGACATCACAGAGCCGGACAGGAGATGCTGAGTGTCAACAGAGGTGTG 2215

Qy  1108 ctcatattccgagaacaccctcgccactaaagccatagaaagagtatatgagactgatt 1167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2216 CTCATATTCCGGAGACCACCTCGCCACTAAAGGCCATAGAAGAGTATATGAGACTGATT 2275

Qy  1168 ggcagaaatacctcaagatgccattgggaggtcatccggctctgtatgaatctgag 1227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2276 GGCCAGAAATACCTCAAGGATGCCATTGGGGACCTGTGTTGCACTCAAGGCTTCCATCACAGCA 2335

Qy  1228 gagaactgtgaagtagacccccatcaagtgcacagcgtccagtctggccagagcaccaggcc 1287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2336 GAGAACTGTGAAGTAGACCCCCATCAAGTGCACAGCGTCCAGTCTCCCAAGGAGTGGCTGCTTT 2395

Qy  1288 aacctgcggatgtgtctggaccctgtgagttggccctgcaaggtgtcaactccccattgcgtgttc 1347
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2396 AACCTGCGGATGTGCTGTGAGTTGGCCCTGCCGTGCAAGGTGGTCACTACGGATGGTGTCCCCGCCAACTG 2455

Qy  1348 ccgaggggagctgaaggaggtgtttgcatcatgcggctgcatggtcagagcgggggcgg 1407
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2456 CCGGCGGGAGCTGAAGGAGGTGTTTGCATGCGGCTGCATGGTGTCCAGCGTGCCGGGGGCCGG 2515

Qy  1408 gaggacattgctgacaggctgatcagcgcctcgctcttcctgcccgccc 1467
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2516 GAGGACATTGCTGACAGGCTGATCAGCGCCTCGCTCTTCCTGCGCGCCCTTCCTGCGCGCCGCC 2575
```

```
Qy  1468 atcatgtcgcccagtctgttggactgatgcgtcaggagtaccccagatggagcagaccctcacga 1527
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2576 ATCATGTCGCCCAGTCTGTTTGGACTGATCGTCAGGAGTACCCCAGATGAGCAGACCCTCACGA 2635

Qy  1528 accctcaccctatcgccaaggtatccagaacctggccaacttttccaagtttaccctca 1587
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2636 ACCCTCACCCTCATCGCCAAGGTTATCCAGAACCTGGCCAACTTTTCCAAGTTTACCTCA 2695

Qy  1588 aaggaggacttcctgggcttcatgaacgagttctggagctggaatggggttctatgcag 1647
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2696 AAGGAGGACTTCCTGGGCTTCATGAACGAGTTTCTGGACTGGAGTGGGGTTCTATGCAG 2755

Qy  1648 caattcttgtatgagatatccaacctgttaacaacctgacacagcacagtttgagggc 1707
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2756 CAATTCTTGTATGAGATATCCAACCTGGACACACTGACCAACAGCACAGTTTGAGGGC 2815

Qy  1708 tacatagacttggccgcgagctctccacacttcacgcctgctctgggaggtgctgccc 1767
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2816 TACATAGACTTGGCCGCGAGCTCTCCACACTTCACGCCCTGCTCTGGGAGGTGCTCCCC 2875

Qy  1768 cagctcagcaaggaagccctcctgaagtctgggccggctgccccggctgccctcagcgcacatc 1827
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2876 CAGCTCAGCAAGGAAGCCCTCCTGAAGTCTGGGCCCGTGCCCGGCTGCCTCCAGCCGCACATC 2935

Qy  1828 agcacagcccgaggaacccaacatccaaaggagcgaggagcccagagcgagcgcgct 1887
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2936 AGCACAGCCCGTGAGGAACCCTAACATCCAAAGGCCAGCCGAGCCGACCAGCGACCGCGCT 2995

Qy  1888 cggtctcagcccatagtgctgcgcgggggccgtcagcgcaaatgcaggcgtacatgatgcgg 1947
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2996 CGGTCTCAGCCCATGGTGCTGCGCGGGGCCCGTCAGCCGCAGATGCCAGGGCTACATGATGCGGG 3055

Qy  1948 gacctcaacagtcatcagacctcagtcctcatggctcggcagctctatg 2007
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  3056 GACCTCAACAGCTCCAACGACCTTCAGTCCTTCAGTGGCTCGAGGCTCTATG 3115

Qy  2008 gacatggctcgctccccctccccccaacaaggagaaaccccccgcccccctcccggtggg 2067
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3116 GACATGGCTCGCTCCCCCTCCCCCCCAACAAGGAGAAACCCCGCCCCGCCCCCTCCCGGTGGG 3175

Qy  2068 ggtaagacctgttctatgtgagccggccaccactggcggccggctcctcccagcatactgc 2127
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3176 GGTAAAGACCTGTTCTATGTCTATGTGAGCCGGCCACCACCTGGCCGGCCGCATACTGC 3235

Qy  2128 acgagcagctcggacatcacagaccggagcagaaagatgctgagtgtcaacaagagtgtg 2187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3236 ACGAGCAGCTCGGACATCACAGAGCCGGAGCAGAGATGCTGAGTGTCAACAGAGGTGTG 3295

Qy  2188 tccatgtggacctgaccgtgttgggacctgtgcactcaagcagcctcactgacacgagtagtgtttcc 2247
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3296 TCCATGTGGACCTGGACCTGTTGGGACCTGTGCACTCAAGCCTTAACAGCAGTAGTGTTTCC 3355

Qy  2248 aacctggcagctgttgggacctgttgcactcaagcaggcttcactgacagcagcccttg 2307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3356 AACCTGGCAGCTGTTGGGGACCTGTGCACTCAAGCCTTAACAGCCAGTAGTGTTTCATCAGCA 3415

Qy  2308 gggttgcggcctgcacctgccggggcctgctcctccaagggagtggctcttcatcacagca 2367
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3416 GGGTTGCGGCCTGCACCTGCCGACCTGCTCCGGGCTCTTCCATCACAGCA 3475

Qy  2368 gccggcatgcgcctcagccagatggtgtcactacggatggtgtccccgccaactg 2427
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3476 GCCGGCATGCGCCCTCCAGCCAGCAGATGGTGCACTACGGATGGTGTCCCCCAGCCAACTG 3535

Qy  2428 cgcatccctcttcctcagaaccctctcttcaatatggctggctgccgatggaccaggcc 2487
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3536 CGCATCCCTCTTCCTTCCTTCCAGAACCCTCCTCTTCCATATGGCTGGCTGCCAGGCCC 3595

Qy  2488 ccagcaggccatggaggagcagtggccatggtccacctcctcccatcaaccaccac 2547
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3596 CCAGCAGGCCATGGAGGAGCAGTGGCCATGGTCCACCTCCTCCCATCCACCACCAC 3655
```

```
Qy   2548  caccatcaccaccgaggggagaaccccaggggacaactttttgcccgttccatggc  2607
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3656  CACCATCACCACCGAGGGGGAGAAACCCCAGGGGACAACTTTTTGCCCGTTCCATGGC  3715

Qy   2608  tatagcaagagacgaggaccctctacagggtccctaagccccctgcggcctccatcctt  2667
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3716  TATAGCAAGAGCGAGGACCTCTCTACAGGGTCCCTAAGCCCCTGCGGCCTCCATCCTT  3775

Qy   2668  cacagccacagctacagtgatgagtttggaccctctggtactgattttaccgtcggcag  2727
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3776  CACAGGCCACAGCTACAGTGATGAGTTTGGACCCTCTGGTACTGATTTTACCGTCGGCAG  3835

Qy   2728  cctcacttcaggacaaaccacatgctctcccgcccagatcaccatcgtccc  2787
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3836  CTTCACTTCAGGACAACCTACAGCACATCGTCTCCCCGCCCAGATCACCATCGGTCCC  3895

Qy   2788  cagaggcagctccctcaggccaggaggggcagtggtggggcagtggtggggggcggt  2847
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3896  CAGAGGGCCAGCTCCCTCAGGGCCAGGAGGGGCAGTGGTGGGGCAGTGGTGGGGGGCCGGT  3955

Qy   2848  ggggccagcaacctccttgagagggcaaatctcagcagttgacagtgagtgctgcc  2907
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   3956  GGGGGCCCAGCCACCTCCCTTGCAGAGGGCAAATCTCAGCAGTTGACAGTGAGTGCTGCC  4015

Qy   2908  cagaaaccccggcgccgtccagcggggaacctattgcagtcccggaaccaagttatggtcct  2967
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4016  CAGAAAACCCCGGCCGTCCAGCGGGAACCTATTGCAGTCCCCGGAACCAAGTTATGGTCCT  4075

Qy   2968  gcccgtccacggcaacagagcctcagcaaagagggcagcattgggggcagcggggcagc  3027
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4076  GCCCGTCCACGGCAACAGAGCCTCAGCAAAGAGGGCAGCATTGGGGGCAGCGGGGCAGC  4135

Qy   3028  ggtggcggaggggggtgggggctcaagcccctccatccaagcagcattcccagactcca  3087
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4136  GGTGGCGGAGGGGGGTGGGGGCTCAAGCCCTCCATCCACCAGCAGCAGCAGCATTCCCAGACTCCA  4195

Qy   3088  tccacgctgaaccccacgatgccggcctcggacgtagcctggttgtccaatatg  3147
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   4196  TCCACGCTGAACCCCACGATGCCGGCCTCGGCGGCCTCCGGAGCCGGACTGTAGCCTGGGTGTCCAATATG  4255

Qy   3148  cctcaccgtccgctgacatcgagagtgcacacattgagcggaaggaagactgaag  3207
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4256  CCTCACCGTCCCCGTGACATCGAGAGTGCACACATTGAGCGGGAAGAAGAGTACAAGCTGAAG  4315

Qy   3208  gagtactgaagtccatggacgagagccgactggacagggtgaaggactacgagaggag  3267
           |||||||||||||||||||||||||||||||||||||||
Db   4316  GAGTACTCGAAGTCCATGGACGAGAGCCGACTGGACAG------GGAGTACGAGGAGGAG  4369

Qy   3268  atccactcactgaagaaaggctacacatgtccaaccggaagtggaagagtacgacgg  3327
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4370  ATCCACTCACTGAAGGAAAGGCTACACATGTCCAACCGGAAGCTGGAAGCGGACAGTACGAGCGG  4429

Qy   3328  aggctgtgtccaggaagaagcagcagcagaatcctgatgcagtcagtaccaagcccgctg  3387
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4430  AGGCTGCGTGTCCCAGGAAGAAGCAGCAGCAGAATCCTGATGCAGTAGCAGTACCAAGCCCGCCTG  4489

Qy   3388  gagagagcgagagcgcttgagcgcagcaggtgagaagaggactcccagatcaagagc  3447
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   4490  GAGCAGAGCGAGAGCGCGAAGCGCTTGAGCGCAGCAGGTGGAGAAGAGGACTCCCAGATCAAGAGC  4549

Qy   3448  atcattgg--------caggctgatgctggtggaggaggagctgcgcggggacca  3494
           |||||||        |||||||||||||||||||||||||||||||||||
Db   4550  ATCATTGGCAGCCGGTCCCTTCAGGCTGATGCTGGTGGAGGAGGAGCTGCGCCGGGACCA  4609

Qy   3495  ccccgccatggctgctgagccgcgtcctgaaccaagagagaggctgctcgacgctca--gagag  3553
           ||||||||||||||||||||||||||||||||||||||||||||||     ||||
Db   4610  CCCCGGCCATGGCTGAGCCGCGTCCTGCAACCCAAGAGAGGCTGCTCGACGCTGCAGGAGAG  4669

Qy   3554  gcagcttcccccttgggtccaacaaaccccgtgtgacgctgccccaccttggaacg  3613
           |||||||||||||||||||||||||||||||||||||||||||||||
Db   4670  GCAGCTTCCCCCTTGGGTCCAACAAACCCCGTGTGACGCTGCCCCACCTTGGAACGG  4729

Qy   3614  cctgccccccagcccacccccaccccccgctgagcccccggctcagccgcgatcaacagagaacgggcgagtt  3673
```

mRNA   /clone="RP4-570F3"
            complement(join(1012..1185,1669..2740,3135..3176,
            3870..4048,4255..4456,4664..4900,6710..6854,7066..7210,
            7338..7961,10016..10114,10324..10477,12824..12945,
            13378..13469,19727..19832,22032..22153,25299..25366))
            /gene="dJ570F3.1"
            /note="match: cDNAs: Em:AF048976 Em:AF058789 Em:AF058790
            Em:AB016962 Em:AF050183 Em:AF047711 Em:AF053938"
            /evidence=not_experimental
            /product="dJ570F3.1 (homolog of the rat synaptic ras
            GTPase-activating protein p135 SynGAP)"

gene   complement(1012..25366)
            /gene="dJ570F3.1"

CDS   complement(join(1012..1185,1669..2740,3135..3176,
            3870..4048,4255..4456,4664..4900,6710..6854,7066..7210,
            7338..7961,10016..10114,10324..10477,12824..12945,
            13378..13469,19727..19832,22032..22153,25299..25320))
            /gene="dJ570F3.1"
            /note="match: proteins: Tr:095174 Tr:044242 Wp:CE23473
            Wp:CE23471 Wp:CE23475 Wp:CE23472 Wp:CE23470 Wp:CE23474
            Wp:CE23476"
            /codon_start=1
            /evidence=not_experimental
            /product="dJ570F3.1 (homolog of the rat synaptic ras
            GTPase-activating protein p135 SynGAP)"
            /protein_id="CAB63778.1"
            /db_xref="GI:6624587"
            /translation="MSYAPFRDVRGPSMHRTQVVHSPYDRPGWNPRFCIISGNQLLML
            DEDEITHPLLIRDRRSESSRNKLRRTVSVPVEGRPHGEHEYHLGRSRRKSVPGGKQYS
            MEGAPAAPRPSQGFLSRRLKSSIKRTKSQPKLDRTSFRQLPRFRSADHDRARLMQ
            SFKESHISHESLLSPSSAAEALELNLDEDSIIKPVHSSILGQEFCFEVTTSSGTKCFAC
            RSAAERDKWIENLQRAVKPNKDNSRRVDNVLKLWIIEARELPPKKRYCELCLDDMLY
            ARTTSKPRSASGDTVFWGHEFEFNNLPAVRALRLHLYRDSDKKRKKDKAGVGLVTVP
            VATLAGRHFTEQWYPVTLPTGSGSGMGSGGGSGGGSGGGKGKGGCPAVRLKARYQ
            TMSILPMELYKEFAEYVTNHYRMLCAVLEPALNVKGKEEVASALVHLLQSTGKAKDFL
            SDMAMSEVDRFMEREHLIFRENTLATKAIEEYMRLIGQKYLKDAIGEFIRALYESEEN
            CEVDPIKCTASSLAEHQANLRMCCELALCKVVNSHCVFPRELKEVFASWRLRCAERGR
            EDIADRLISASLFLRFLCPAIMSPSLFGLMQEYPDEQTSRTLTLIAKVIQNLANFSKF
            TSKEDFLGFMNEFLELEWGSMQQFLYEISNLDTLTNSSSFEGYIDLGRELSTLHALLW
            EVLPQLSKEALLKLGPLPRLLNDISTALRNPNIQRQPSRQSERPRPQPVVLRGPSAEM
            QGYMMRDLNSSIDLQSFMARGLNSSMDMARLPSTPTKEKPPPPPGGSKDLFYVSRPPL
            ARSSPAYCTSSSDITEPEQKMLSVNKSVSMLDLQGDGPGGRLNSSSVNLAAVGDLLH
            SSQASLTAALGLRPAPAGRLSQGSGSSITAAGMRLSQMGVTTDGVPAQQLRIPLSFQN
            PLFHMAADGPGPGPDGGGHGPPSSHHHHHHHRGGEPGDTFAPFHGYSKSED
            LSSGVPKPRAASILHSHSYSDERGPSGTDFTRRQLSLQDNLQHMLSPPQITIGPQRPA
            PSGPGGSGGGSGGGGGQPPPLQRGKSQQLIVSAAQKPRFSSGNLLQSPEPSYGPAR
            PRQQSLSKEGSIGGSGSGGGGGLKPSITKQHSQTPSTLNPTMPASERTVAWVSNM
            PHLSADIESAHIEREEYKLKEYSKSMDESRLDR"

repeat_region   2178..2210
                /note="11 copies 3 mer ggt 90 conserved"

repeat_region   2945..3122
                /note="89 copies 2 mer gg 56 conserved"

misc_feature   complement(4211..4599)
                /gene="dJ570F3.1"
                /note="match: GSS: Em:B68206"

repeat_region   4484..4539
                /note="28 copies 2 mer ca 94 conserved"

repeat_region   4605..4649
                /note="3 copies 15 mer 91 conserved"

repeat_region   4903..4970
                /note="34 copies 2 mer ga 70 conserved"

misc_feature   4996..5190
                complement(6393..6846)
                /gene="dJ570F3.1"

repeat_region   8077..8178
                /note="13 copies 15 mer 59 conserved"
                /note="match: GSS: Em:AQ475907"

repeat_region   8476..8841
                /note="MLT1C repeat: matches 6..109 of consensus"

repeat_region   10989..11119
                /note="MLT1C repeat: matches 109..461 of consensus"

repeat_region   13130..13161
                /note="FLAM_C repeat: matches 3..131 of consensus"

---

repeat_region   13256..13303
                /note="16 copies 2 mer ca 93 conserved"

misc_feature   17111..17815
                /note="24 copies 2 mer ct 75 conserved"
                /note="CpG island"
                /evidence=not_experimental

repeat_region   17369..17458
                /note="6 copies 15 mer 70 conserved"

repeat_region   17370..17513
                /note="72 copies 2 mer cc 63 conserved"

repeat_region   17515..17658
                /note="72 copies 2 mer gg 59 conserved"

repeat_region   17631..17726
                /note="32 copies 3 mer gcc 63 conserved"

repeat_region   19005..19078
                /note="37 copies 2 mer ga 67 conserved"

repeat_region   19612..19691
                /note="40 copies 2 mer cc 63 conserved"

repeat_region   19958..19911
                /note="27 copies 2 mer cc 70 conserved"

repeat_region   24653..24808
                /note="78 copies 2 mer cc 57 conserved"

repeat_region   24706..24831
                /note="14 copies 9 mer ctcctccct 61 conserved"

repeat_region   25104..25289
                /note="93 copies 2 mer cc 59 conserved"

repeat_region   25187..25294
                /note="12 copies 9 mer ctcccccc 64 conserved"

repeat_region   25367..25465
                /note="11 copies 9 mer gggagag 69 conserved"

repeat_region   25373..25522
                /note="10 copies 15 mer 64 conserved"

repeat_region   25413..25469
                /note="19 copies 3 mer agg 86 conserved"

misc_feature   26820..27191
                /note="match: GSS: Em:AQ605113"

misc_feature   27145..27728
                /note="CpG island"
                /evidence=not_experimental
                complement(27412..27910)

misc_feature   join(27713..28149,28320..28379,28488..28533,28661..28710,
                28854..29088)
                /gene="CICK07210.5"
                /note="match: cDNAs: Em:AF106943 Em:E06811
                match: ESTs: Em:W75983 Em:AA187766 Em:AI004315 Em:R77008
                Em:N94792 Em:AA843316 Em:AA384803 Em:AI245618 Em:AI921804
                Em:AI792323 Em:AI802586 Em:AI524460 Em:AA843230
                Em:AA887246 Em:AI096824 Em:AI929640 Em:N32014 Em:AA725127
                Em:AA357305 Em:AF034177 Em:W39055 Em:AA622278 Em:AA069440
                Em:AI087362 Em:AI050991 Em:AI340323 Em:AI028045
                Em:AA934024 Em:AI073391 Em:AA331045 Em:AA708233
                Em:AI342398 Em:AA056366 Em:AI582796 Em:AI762297
                Em:AI340198 Em:AI541387 Em:N34313 Em:AI567726 Em:AI364919
                Em:AI310946 Em:AA181976 Em:AI042256 Em:881592 Em:T49588
                Em:AI929571 Em:D52353 Em:AA035663 Em:N48534 Em:AI373814
                Em:AA971499 Em:H25294 Em:AA703121 Em:AA446226 Em:AA379516
                Em:AI313425 Em:W70287 Em:AA937820 Em:H21228 Em:AA471384
                Em:N71594 Em:AA095718 Em:AA377858 Em:H25235 Em:H52756
                Em:D52710 Em:AA329903"
                /evidence=not_experimental
                /product="dJ570F3.2 (POLYPEPTIDE FROM PATENTED CDNA
                EMBL:E06811)"

mRNA   Query Match   23.1%;   Score 920.8; DB 69; Length 94770;
                Best Local Similarity 91.0%;   Pred. No. 2.2e-167;
                Matches 979; Conservative 0; Mismatches 97; Indels 0; Gaps 0:

                Qy 1997 acagctctatggacatggctcgcctccctcccaaccaaggagaaacccccgccgcccc 2056
                         ||||||||||||||||||||||||||||||||||||||| |||||| || |||
                Db 2744 ACAGCTCTATGGACATGGCTCGCCTCCCTCCCAACCAAGGAAAAGCCACCCCCACCAC 2685

```
ACCESSION    AL161903
VERSION      AL161903.12  GI:10039702
KEYWORDS     HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 150956)
  AUTHORS    Tracey,A.
  TITLE      Direct Submission
  JOURNAL    Submitted (28-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
             CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
             requests: clonerequest@sanger.ac.uk
COMMENT      On Sep 9, 2000 this sequence version replaced gi:9996876.
             ----------------- Genome Center
             Center: Sanger Centre
             Center code: SC
             Web site: http://www.sanger.ac.uk
             Contact: humquery@sanger.ac.uk
             -------------- Project Information
             Center project name: bA175A4
             ---------------- Summary Statistics
             Assembly program: XGAP4; version 4.5
             Sequencing vector: plasmid; L08752; 100% of reads
             Chemistry: Dye-terminator Big Dye; 99% of reads
             Chemistry: Dye-primer-amersham; 0% of reads
             Chemistry: Dye-primer Big Dye; 0% of reads
             Consensus quality: 145680 bases at least Q40
             Consensus quality: 147560 bases at least Q30
             Consensus quality: 148600 bases at least Q20
             Insert size: 150756; sum-of-contigs
             Insert size: 148326; 10.8% error; agarose-fp
             Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality
             coverage: 6.54x in Q20 bases; agarose-fp
             -------------------
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 3 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *       1  54611: contig of  54611 bp in length
             *   54612  54711: gap of      100 bp
             *   54712  70293: contig of  15582 bp in length
             *   70294  70393: gap of      100 bp
             *   70394 150956: contig of  80563 bp in length.
FEATURES             Location/Qualifiers
     source          1..150956
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-175A4"
                     /clone_lib="RPCI-11.1"
     misc_feature    1..54611
                     /note="assembly_fragment:01454
                     clone_end:T7
                     vector_side:left"
     misc_feature    54712..70293
                     /note="assembly_fragment:02542
                     fragment_chain:1"
     misc_feature    70394..150956
                     /note="assembly_fragment:02844
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:right"
BASE COUNT   39575 a  37381 c  36411 g  37389 t  200 others
ORIGIN
```

```
Qy   2057 ctcccgtggggtaaagacctgttctatgtgagccggccaccactggcccggtcctccc 2116
          | | ||||  |||||||||||||||||||||||||||| |||| |||| ||||||  |
Db   2684 CGCCTGGTGGTGGTAAAGACCTGTTCTATGTAAGCCGTGCCACCCCTGGCCCGTTCCTCAC 2625

Qy   2117 cagcatactgcacgagcagctcggacatcacagagcggagcagaagatgctgagtgtca 2176
          ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2624 CAGCATACTGCACGAGCAGCTCGGACATCACAGAGCAGAAGATGCTGAGTGTCA 2565

Qy   2177 acaagagtgtccatgctggacctgcaggcgcacggggcctggggcgccttaacagca 2236
          ||||||||||||||||||| | ||||||| || ||||||||||||||| ||||||||
Db   2564 ACAAGAGTGTGTCCATGCTGGACTTACAGGGTGATGGGCTGTGGTGGCCGCCTCAACAGCA 2505

Qy   2237 gtagtgtttccaacctggcagctgttgggggacctgttgcactcaagccaggcttcactga 2296
          |||| |||| ||||||| || |||  ||||||||||||||  || ||| || |||||
Db   2504 GCAGTGTTTCGAACCTGGCCGCGCCGTAGGGGACCTTGCTGCTCGCACTCAAGCCAGGCCTCGCTGA 2445

Qy   2297 cagcagccttggggtttgcggctcggacctgcccggcgcctctccaaggggatggctctt 2356
          ||||||||||||| ||||||| ||| |||||| |||||||||||||||||||  ||||||
Db   2444 CAGCAGCCTTGGGGCTTACGGGCCTGCCCGACGGGCCTCTCCCAGGGAGTGGCTCAT 2385

Qy   2357 ccatcacagcagcacggcatgcgcctcagccagatggtgtcactacggatggtgtccccg 2416
          |||||| ||  |||||||||||||||||||||||||| ||| || ||||||||||| | |
Db   2384 CCATCACGGCGGCCTGCGCCTGCGCATGCGGTCAGCAGCATGGGTCACCACACAGCGGTGTCCCTG 2325

Qy   2417 cccagcaactgcgcatcccctctttcctttccagaaccctctcttccatatggctgccgatg 2476
          ||||||||||||| ||||  || ||||| ||| ||||| |||| ||||||||||||||||
Db   2324 CCCAGCAACTGCGAATTCCCCTCTCCTTCCAGAACCCTCTCTTCCACATGGCTGCTGATG 2265

Qy   2477 gaccagggcccccagcaggccatggaggagcagtggccatggtccacctttcctcccatc 2536
          | |||| |||||| ||||||||||||| |||||||| ||| |||| |||||||||||||
Db   2264 GGCCAGGTCCCCCAGGCGGCCATGGAGGAGGGGCGGTGGCCATGGCCGTGGCCACCTTCCTCCCATC 2205

Qy   2537 acaacaccaccacacacatcaccatcaccgaggggaagaaccccaggggacacctttgccc 2596
          ||| |||||||||||||||||||||||||||||||||||| |||||||  || ||||||
Db   2204 ACCACCACCACCATCACCACCATCACCGAGGGTGGAGAGCCCGTGGGACACCTTTGCCC 2145

Qy   2597 cgttccatggctatagcaagagggaggagaccctctcacagggtccctaagcccccctcgg 2656
          | |||||||||| |||||||||||||||||||||| |||||||| |||||||||||||||
Db   2144 CATTCCATGGCTATAGCAAGAGTGAGGACCCTTCTCCGGGGTCCCCAAGCCCCCCTGCTG 2085

Qy   2657 ctccatccttcacacaccacagctacagtgatgagttggaccctctggtactgatttta 2716
          ||||||||||||||||||||||||| | |||||||| ||||||| ||||  ||||||||
Db   2084 CCTCCATCCTTCACAGCCACAGCTACAGTGATGATGAGTTTGGACCCTCTGCCACTGACTTCA 2025

Qy   2717 ccgtgcagcagctcacttcaggacaactacagcacatgctctccccgcccccagatca 2776
          |||||||||| ||||||||||||||||||||||||||||| ||||||||||||||||||
Db   2024 CCCGTCGGCCAGCTTTTCACTCCAGGACAACCTGCACCCTGGCGTCTCCCCTCCCCAGATCA 1965

Qy   2777 ccatcggtccccagagggcagctccctccagggcagtggtggtgggggcagtg 2836
          |||||||||||||||||||||| |||||||||||||||||| |||| | | |||||||
Db   1964 CCAATTGGCGTCCCCAGAGGGCCAGCAGCCCTGGAGGTGGACAGCGCCGGTGGGGGCAGCG 1905

Qy   2837 gtggggccggtggggccggtggggccaccctcagttgacag 2896
          ||| | ||||||||||||||||||||||| ||||||||||||
Db   1904 GTGGGGGTGGCGGGGGCCAGCCCAGCGGGCCAGTCTCAGCAGTTGACAG 1845

Qy   2897 gttatggtcctgccccgtccctccacggcaacagagcctcagcaaagaggggcagcattggggca 2956
          || || ||||| ||||||||||||||||||||||| ||||| |  |||||  |||
Db   1844 TCAGCGACCGCCCAGAGAAACCCCGCCCGTCCACGGCAACAGGGGCAGCATTGGGGCA 1785

Qy   3017 gcggggcagcggtggcgcggggggttggggggctcaagcccctccatcaccaagcag 3072
          |||||||||||| ||  | ||||||||||||||||||||| |||||||||||||||||
Db   1724 GCGGGGCAGCCAGCCGGTGGCCGGAGGGGTGGGGGGCTGCAAGCCCTCCATCACCAAGCAG 1669
```

RESULT 7
AL161903/c
LOCUS   AL161903 150956 bp   DNA   HTG   29-SEP-2000
DEFINITION Homo sapiens chromosome 6 clone RP11-175A4, *** SEQUENCING IN

Query Match   23.1%;   Score 920.8;   DB 70;   Length 150956;

Best Local Similarity 91.0%; Pred. No. 1.9e-167;
Matches 979; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```
Qy    1997  acagctctatggacatggctcgctcccctcccaccaagagaaccccgcgcgccc  2056
            ||||||||||||||||||||| |||||||||||| || ||| |||||  ||
Db   99893  ACAGCTCCATGGACATGGCTCGCTCGCCTCCCCTCCCCAACCAAGGAAAAGCCACCCCCACCAC  99834

Qy    2057  ctcccggtggggtaaagacctgttctatgtgagccggccaccactggccggtcctccc  2116
            |||||||||||| |||||||| || |||||||||||||||| ||| |||||||||||
Db   99833  CGGCTGGTGGTGTGAAAGACCTGTTCTATGTAAGCCGTCCACCCCTGGCCCGGTTCCTCAC  99774

Qy    2117  cagcatactgcacgagcagctcggacatcacagagcggagcagagatgctgagtgtca  2176
            |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   99773  CAGCATACTGCACGAGCAGCTCGGACATCACAGAGCCAGAGCAGAAGATGCTGAGTGTCA  99714

Qy    2177  acaagagtgtccatgctggacctgcagggcgacgggcgggcgccttaacagca  2236
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   99713  ACAAGAGTGTGTCCATGCTGGACTTACAGGGTGATGGGCCGCCTGGTGCCGCCTCAACAGCA  99654

Qy    2237  gtagtgtttccaacctgcagctgttggggaccctgtgcactcaagccaggcttcactga  2296
            |||||||| |||||||||| ||||||||||| || |||||||||||||| || ||
Db   99653  GCAGTGTTTCGAACCTGCGGCCGTAGGGACCTGCTGCACTCAAGCCAGGCCCTCGCTGA  99594

Qy    2297  cagcagccttggggttgcggcctgcacctgccgggcgcctctcccaagggagtggctcctt  2356
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99593  CAGCAGCCTTGGGGCTACGGGCCTGCGCCTGCCGGACGCCTCTCCCAGGGAGTGGCTCAT  99534

Qy    2357  ccatcacagcagccggcatgcgcctcagccagatgggtgtcactacggatggtgtcccg  2416
            |||||| ||||||||||||||||||||||||||||||||||| ||| ||||||||
Db   99533  CCATCACGGCGGCCGGCATGCGCCTGGCCTCCAGCCAGATGGGTGTCACCACAGACGGTGTCCCTG  99474

Qy    2417  cccagcaactgcgcatccctcttcttccagaaccctcttccatatggctgcccgatg  2476
            |||||| ||||||||||||||||||||||||||||||| ||||||||||||||||
Db   99473  CCCAGCAACTGCCGAATCCCCTCCTTCCAGAACCCTCTTCCACATGGCTGCTCGTGATG  99414

Qy    2477  gaccagggccccagcaggccatgggaggagcagtgccatgtccacctcctcccatc  2536
            | ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   99413  GGCCAGGTCCCCAGGCGGCCATGCGAGGGCGGTGGCCATGGCCACCTTCCTCCCATC  99354

Qy    2537  acaccacaccaccatcaccacaccaccaccgagggggaaaccccaggggacacctttgccc  2596
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99353  ACCACCACCACCACCATCACCACCACCACCGAGGTGACGGTGGCAGAGCCCCCTGGGGACACCTTGCCC  99294

Qy    2597  cgttccatggctatagcaaggacgaggacctctctcacagggtcctaagccccctgcgg  2656
            |||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   99293  CATTCCATGGCTATAGCAAGGAGGACCTCTCTCCCGGGGTCCTAAGCCCCCTGTG  99234

Qy    2657  ctccatccttcacagccacagctacagtgatgagtttggaccctctgtactgatttta  2716
            |||||||||||||||||||||||||||| ||| ||||||| ||||||| || |||
Db   99233  CCTCCATCCTTCATAGTCCATGATGAGTTTGGACCCCTTCTGGCACTTGACTTCA  99174

Qy    2717  ccgtcggcagctctcacttcaggacaacctacagcacatgtctcccccgcccagatca  2776
            |||||||||||||| ||||| |||||||||||||| ||||||||||||||||||
Db   99173  CCCGTCGGCAGCTTTCACTCCAGGACAACCTGCAGCACATGTGTCCCCTCCCCCAGATCA  99114

Qy    2777  ccatcggtcccagaggccagctccctcagggccaggaggggcagtggtggggggcagtg  2836
            |||| ||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   99113  CCATTGGTCTCCCCAGAGGCCCTGGAGGTGGAGCGGTGGGGGGCGTGGGGGGGCATGGG  99054

Qy    2837  gtggggcggtgggggccagcacctcaaatctcagcagttgacag  2896
            ||||| ||||||||||||||||||||||||||||||||||| ||
Db   99053  GTGGGGGCGGTGGGGGCGGGGGCAGCGCCTCCATTGCAGAGGGGCAAGTCTCAGCAGTTGACAG  98994

Qy    2897  tgagtgtgccccagaaaaccccgccgtccagcgggaaacctattgcagtccccggaaccaa  2956
            |||| ||||||||||||||| |||||||||||||||||| |||||||| |||||
Db   98993  TCAGCGCAGCCCAGCCAGAAAACCCCGGCCATCCAGCGGGAATCTATTGCAGTCCCCAGAGCCAA  98934

Qy    2957  gttatggtcctgcccgtccacgcaacagagggcagcattgggggca  3016
            |||||||||||||| |||||||||||||||||||||||||||| ||
Db   98933  GTTTATGGCCCCGCCCCGTCCACGGCCACGCAACAGAGCCTCAGCCAAGGAGGAGCCAGCATTGGGGGCA  98874

Qy    3017  gcggggcagcggtggcggagggggtggggcgtcaagcctccatcaccaagcag  3072
            |||||||||||||||||||||||||||||||| |||||||||||||||||||
```

```
*     80765     83307: contig of 2543 bp in length
*     83308     83407: gap of      100 bp
*     83408     88383: contig of 4976 bp in length
*     88384     88483: gap of      100 bp
*     88484     93697: contig of 5214 bp in length
*     93698     93797: gap of      100 bp
*     93798     99354: contig of 5557 bp in length
*     99355     99454: gap of      100 bp
*     99455    111055: contig of 11601 bp in length
*    111056    111155: gap of      100 bp
*    111156    118645: contig of 7490 bp in length
*    118646    118745: gap of      100 bp
*    118746    122421: contig of 3676 bp in length
*    122422    122521: gap of      100 bp
*    122522    132468: contig of 9947 bp in length
*    132469    132568: gap of      100 bp
*    132569    137959: contig of 5391 bp in length
*    137960    138059: gap of      100 bp
*    138060    151691: contig of 13632 bp in length
*    151692    151791: gap of      100 bp
*    151792    158253: contig of 6462 bp in length
*    158254    158353: gap of      100 bp
*    158354    169323: contig of 10970 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..169323
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone_lib="RP11-567N9"
                     /clone_lib="RPCI-11.2"
     misc_feature    1..5456
                     /note="assembly_fragment:00884
                     fragment_chain:1"
     misc_feature    5557..8767
                     /note="assembly_fragment:00832
                     fragment_chain:1"
     misc_feature    8868..17168
                     /note="assembly_fragment:01060
                     fragment_chain:1"
     misc_feature    17269..26591
                     /note="assembly_fragment:01172
                     fragment_chain:1"
     misc_feature    26692..31188
                     /note="assembly_fragment:00279
                     fragment_chain:1"
     misc_feature    31289..37579
                     /note="assembly_fragment:01205
                     fragment_chain:1"
     misc_feature    37680..45350
                     /note="assembly_fragment:01334
                     fragment_chain:2"
     misc_feature    45451..48354
                     /note="assembly_fragment:01263
                     fragment_chain:2"
     misc_feature    48455..63401
                     /note="assembly_fragment:00990
                     fragment_chain:2"
     misc_feature    63502..67724
                     /note="assembly_fragment:00045
                     fragment_chain:2"
     misc_feature    67825..70550
                     /note="assembly_fragment:00965
                     fragment_chain:2"
     misc_feature    70651..80664
                     /note="assembly_fragment:00637
                     fragment_chain:3"
     misc_feature    80765..83307
                     /note="assembly_fragment:01291
                     fragment_chain:3"
     misc_feature    83408..88383
                     /note="assembly_fragment:01317
                     fragment_chain:3"
     misc_feature    88484..93697
```

```
     misc_feature    93798..99354
                     /note="assembly_fragment:01171
                     fragment_chain:4"
     misc_feature    99455..111055
                     /note="assembly_fragment:00288
                     fragment_chain:4"
     misc_feature    111156..118645
                     /note="assembly_fragment:00286
                     fragment_chain:4"
     misc_feature    118746..122421
                     /note="assembly_fragment:00303
                     fragment_chain:5"
     misc_feature    122522..132468
                     /note="assembly_fragment:01184
                     fragment_chain:5"
     misc_feature    132569..137959
                     /note="assembly_fragment:00501
                     fragment_chain:6"
     misc_feature    138060..151691
                     /note="assembly_fragment:01244
                     fragment_chain:6"
     misc_feature    151792..158253
                     /note="assembly_fragment:00032"
     misc_feature    158354..169323
                     /note="assembly_fragment:00779"
                     /note="assembly_fragment:01270"
BASE COUNT    41674 a   41332 c   41269 g   42673 t   2375 others
ORIGIN
```

```
Query Match             23.1%;   Score 920.8;  DB 70;   Length 169323;
Best Local Similarity   91.0%;   Pred. NO. 1.9e-167;
Matches  979;   Conservative  0;   Mismatches  97;   Indels  0;   Gaps  0;

Qy   1997  acagctctatggacatggctcgcctccctcccaaccaaggaggaaaccccgccgccgcccc  2056
           ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  83616  ACAGCTCTATGGACATGGCTCGCCTCCCTCCCAACCAAGGAGAAAAGCCACCCCCCACCAC  83675

Qy   2057  ctccggtggggtaaagacctgttctatgtgagccgccaccactggccggtcctccc  2116
           |||||| |||| ||||||||||| |||| |||||||||||| |||||| |||||||
Db  83676  CGCCTGGTGTGGTAAAGACCTGTTCTATGTAAGCCGTCCACCCCTGGCCCGTTCCTCAC  83735

Qy   2117  cagcatactgcacgagcagctcggacatcacagagccggagcagaagatgctgagtgtca  2176
           ||||||||||||||||||||||||| |||| ||||||||||||  ||||||||||||||
Db  83736  CAGCATACTGCACGAGCAGCTCGGACATCACAGAGCCAGAGCAGAAGAATGCTGAGTGTCA  83795

Qy   2177  acaagagtgtgtccatgtggacctggacggcaggcgacggggcctgggggccgcctaacagca  2236
           |||||||||||||||||| | || | |||||| | |||||||| || |||||||||||||
Db  83796  ACAAGAGTGTGTCCATGTGGACTTACAGGGTGATGGGCCTTGGTGGCCGCCTCAACAGCA  83855

Qy   2237  gtagtgtttccaacctggcagctgttgggacctgttcactcaagccaggcttcactga  2296
           | ||||||| |||||||| || |||||| ||||||||||||||||||||| || |||
Db  83856  GCAGTGTTTCGAACCTGGCGCCGTAGGGACCTGCGCACTCAAGCCAGGCCTCAGCCCTGA  83915

Qy   2297  ccatcacagcagccggcatgccctcagccagatgggtgtactacggatggtgtccccg  2356
           ||||||||||||||||||||||||| |||||||||||||||||||||||||| |||||
Db  83916  CCATCACAGCAGCCGGCTGGCATGCCCTCAGCCAGATGGTGTCACCAGACGGTGTCCCTG  83975

Qy   2357  ccagcaactggcatccctcttcctccagaacccttccatatggctgccgatg  2416
           |||||||||||| ||||| | ||||||||||||||||||||||||||||||
Db  83976  CCATCACGGCGGGCTGGCATGCCTCAGCCAGGTGTCACATGGTCTGCGCTGATG  84035

Qy   2417  cccgcaactggcatccctcttcctccagaaccctctcctccatatggctgccgatg  2476
           ||||||||||||||||| || |||||||||||||||||||||||||||||||||||
Db  84036  CCCGCAACTGCGAATCCCCCTCCTTCCAGAACCCTCTCTTCCACATGGCTGCTGATG  84095

Qy   2477  gaccagggccccagcagccagccatggagggacgagtgccatgtccacctcctccatc  2536
           |||||||||| ||||||||||| |||||||||||| ||| |||||||||||||||||||
Db  84096  GGCCAGGTCCCCCAGGCGGCCATGGAGGGGCGGTGGCCGTGGCATGGCCCATGGCCCATC  84155

Qy   2537  accaccaccaccaccatcaccgagggggagaaccccagggggacactttgccc  2596
           ||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  84156  ACCACCACCACCACCATCACCACCACCACCGAGGTGGAGACGCCCCTGGGGACACACTTTGCCC  84215
```

```
            KDNCRRAENVLRLWIIEAKDLAPKKKYFCELCLDDTLFARTTSKTKADNIFWGEHFEF
            FSLPPLHSITVHIYKDVEKKKKDKNNYVGLVNIPTASVTGRQFVEKWPVSTPTPNK
            GKTGGPSIRIKSRFQTITILPMEQYKEFAEFVSYNTMLCYSVLEPVISVRNKEELACA
            LVHILQSTGRAKDFLIDLVMSEVDRCGEHDVLIFRENTIATKSIEEYLKLVGGQYLHD
            ALGEFIKALYESDENCEVDPSKCSSSELIDHQSNLKMCCPLAFCKIINSYCVFPRELK
      2657  EVPASWKQQCLNRGKQDISERLISASLFLRFICPAIMSPSLFNLMQEYPDDRTSRTLT
            LIAKVIQNLANFAKFGNKEEYMAFMNDFLEHEWGGMKRFLEISNPDTISNTPGFDGY
            IDLGRELSVLHSLLWEVVSQLDKGENSFLQATVAKLIGPLPRVLADITKSLTNPPIQQ
            QLRRFTEHNSSPNVSGSLSSGLQKTFEDPTDSDLHKLKSPSQDNTDSYFRGKTLLLVQ
      2717  QASSQSMTYSEKDERESSLPNGRSVSLMDLQDTHAAQVEHASVMDLVPFRLTGSQLSI
            TQVASIKQLRETQSTPQSAPQVRRPLHPALNQPGGLQPLSFQNPVYHLNNPIPAMPKA
            SIDSSLENLSTASSRSQSNSEDFKLSGPSNSSMEDFTKRSTQSEDFSRRHTVPDRHIP
            LALPRQNSTGQAQIRKVDQGLGARAKAPPSLPHSASLRSTGSMSVVSAALVAEPVQN
      2777  GSSRSRQQSSSRESPSVPKVRAIQROQTQQVQSPVDSATMSPVERTAAWVLNNGQYEED
            VEETEQNLDEAKHAEKYEQEITKLKERLRVSSRRLEEYERRLLVQEOQMOKLLEYKA
            RLEDSEERLRRQQEERDSQMKSIISRLMAVEELKKDHAEMQAVIDAKOKIIDAQEKR
      2837  IVSLDSANTRLMSALTQVKERYSMQVRNGISPTNPYKLSITENGEFKNSSC"

BASE COUNT      1292 a      1022 c      1012 g      1042 t
ORIGIN
```

```
Query Match             15.8%;    Score 629.4;  DB 10;  Length 4368;
Best Local Similarity   62.5%;    Pred. No. 4.3e-111;
Matches 1080;  Conservative   0;  Mismatches 571;  Indels   78;  Gaps   3;

Qy      52  caaggcttcctgagccggaggtaaaaagctcatcaaacgtacaaagtcacaacccaa  111
            ||||| ||||| | ||| |||||   |||||||| || ||  || | ||   ||
Db     243  CCAGGATTCTTCAGCAAGGCCGTGAAGGCTCCATCAAGAGGCTCCAAAAGCCAGTCAAAG  302

Qy     112  cttgaccggaccagcagctttgacagatctggctctgctcccgaagtgctgaccatgac  171
            |||||||||| |  ||||||||||| |               ||| | ||| |||
Db     303  CTTGACAGAAACACGAGCTTTCG-------GCTTCCATCCTTCGCAGTACAGATGAC  353

Qy     172  cgggccggctgatgcagagcttcaaggagtctcactcccatgagtccctgctggatccc  231
            ||||| | ||| | || |||| |||| |||||| ||  || |  |||||||| ||
Db     354  AGGTCTCGTGGGCGTGCCTAAACTAAAAGAGTCACGTTCCCATGAATCCTTGCTGAGCCCA  413

Qy     232  agcagtgctgtgagggcctggagctcaacctggatgaagactccattacaaggcagta  291
            | ||||||| ||| |||   || | ||||  |  ||| |||| |||||| | ||||
Db     414  TGCAGCACAGTGGAATGTCTGGATCTTGGTAGAGGGGAACCTGTATCAGTGAAACCACTT  473

Qy     292  cacagtccatcctgggccaggagttctgctttgaggtaacaacatcgtctgggacaaaa  351
            || ||||||||||||| ||||||| |||||||||||||| ||||| ||||||  ||
Db     474  CATAGTAGCATCCTTGGACAAGACTTCTGCTTTGAGGTTACCTACTTAAGTGGAAGTAAA  533

Qy     352  tgtttgcctgcggtctgcagccgaagggacaaatggattgagaatctacagagggct  411
            ||  || || | ||||||| ||    ||||||| || || |  || |||| ||
Db     534  TGCTTCAGCTGTTAATTCTGCTTCTGAGAGGACAAGTGGATGAAGAAACCTTCGCAGGACA  593

Qy     412  gtgaaacccaacaaggacaacagcgccgggtagataacgtgctgaaactatggatcata  471
            ||  |||| |||||||  || ||||| ||||| ||| |||| |||| |||||
Db     594  GTTCAACCTAATAAGGACAATTGCAGGCGCGAGCTGAAAATGTTCTTCGTTTATGGATCATT  653

Qy     472  gaactcgagagagctgcccccaagaaagcgatattactgcggttatgccggacacatg  531
            |||| |||||| ||| | ||||||||||  |  ||| || || ||||| || |||| |
Db     654  GAAGCCAAGGACCTTGCCCCTAAAAAGAAATATTTCTGCGAACTGTGCCTGATGATACC  713

Qy     532  ctctatgcacgagccgaccacttccagccctccaggagacactgtctttggggc  591
            |||| || | |||||||| |||||||||| ||||||       |||| |||||
Db     714  CTCTTTGCTCGTACAACCAGCAGCAAGACCAA--------AGCAGACAATATTTTCTGGGGC  764

Qy     592  gagcacttgagtttaacaacctgctgctgtccgggcgtgcggctgcatctgtaccgt  651
            |  ||||| ||||||||||||||||||||| || |||||||| ||||  |||||| ||
Db     765  GAACATTTTGAATTCTTCAGCCTTCCACCTCTTCATAGTATCACAGTTCACATTTACAAG  824

Qy     652  gactcggacaaaaagcggaagaaggacaaggcaggctacgttggcctggtgactgttcca  711
            || ||||| |||||| |||| ||||| |||||| | |||| || ||| | |||
Db     825  GATGTGGAAAAAAAGAAAAAAGAAGAACAAGAATAATTATGTAGGGCTAGTCAACATCCCC  884

Qy     712  gtggcaccctggctgggcgccacttcacagacagtggtaccccgtgacccctgcaaca  771
            ||||| || ||| |||||||| | | |||||| ||||| | || || |||
Db     885  ACTGCCAGTGTGACTGGTCGCCAATTTGTAGAAAAGTGGTATCCAGTGAGACCACCTACA  944
```

```
Qy    2597  cgttccatggctatagcaagagcgaggacctctcacaggggtccctaagcccctgcgg   2656
            | ||||||||||||| ||||| ||||||||||||||| |||||||| | |||||||
Db   84216  CATTCCATGGCTATAGCAAGAGTGAGGACCTCTCTCCCGGGGTCCCAAGCCCCCTGCTG   84275

Qy    2657  ctccatcctcacagccacagctacagtgatgagttggaccctctggtactgatttta   2716
            |||||||||||||||||| ||||| |||||||||||||||| ||  |||||| |  |
Db   84276  CCTCCATCCTTCATAGCCACAGCTACAGTGATGAGTTGGACCCTCTGGCACTGACTTCA   84335

Qy    2717  ccgtggcagctcacttcaggacaacctacagacacatgctctccccgccccagatca   2776
            |||||||| |||||||||||| |||||| |  ||||| ||| |||||| ||||||
Db   84336  CCCGTCGGCAGCTTTCACTCCAGGACAACCTGCAGCACAATGCTGTCCCCTCCCCAGATCA   84395

Qy    2777  catcggtccccagagagccagctccctcagggccaggaggggcagtggtggggcagtg   2836
            | ||||| |||||||||||| |||| ||||| ||||| || |||| || |||
Db   84396  CCATTGGTCCCCAGAGAGCCAGCCCCCTCAGGGCCTCAGGGCTGGGCGTGGGGCAGCG   84455

Qy    2837  gtggggcggtgggggccagccagccacctccttgcagagaggggcaaatctcagcagttgacag   2896
            ||||||||||||| ||||| |||| || |||  ||||||||||||||||||| |||| ||||||
Db   84456  GTGGGGGTGGCGGGGGGCAGCCAGCCACGCGCCCTCCATTGCAGAGGGCAACTCTCAGCAGTTGACAG   84515

Qy    2897  tgagtgctgccagaaaccccgggcgtccagcgggaacctattgcagtccccggaaccaa   2956
            | ||||| ||||||| |||||||||| ||||| |||| | ||||| ||||| |||| ||
Db   84516  TCAGCGCCAGCCCCAGAAACCCCGGCCATCCAGCGGGAATCTATTGCAGTCCCCAGAGCCAA   84575

Qy    2957  gttatggtcctgccgtccacggcaacagagctcagcaaaagagaggcagcattggggca   3016
            | |||  | |||| || ||||||||||| |||||  | ||||||||||||| |||||||
Db   84576  GTTATGGCCGCCCCCCGTCCACGGCAACAGAGCTCGGCAAGAGGAGGGCCATTGGGGGCA   84635

Qy    3017  gcgggggcagcgcgtgcgcggaggggtgggggtcaaggcccctccatcaccaaggcag   3072
            |||||||||||||||||| || || ||||||| ||||||||||||||||||||||||||
Db   84636  GCGGGGCAGCCGCGTGCCGCGAGGGGTGGGGGGCTGGGGGGCTGAAGCCCTCCATCACCAAGGCAG   84691
```

```
RESULT       9
AF047711
LOCUS        AF047711        4368 bp        mRNA        PRI        14-JAN-1999
DEFINITION   Homo sapiens nGAP mRNA, complete cds.
ACCESSION    AF047711
VERSION      AF047711.1   GI:4105588
KEYWORDS     .
SOURCE       human.
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 4368)
AUTHORS      Noto,S., Maeda,T., Hattori,S., Inazawa,J., Imamura,M., Asaka,M. and
             Hatakeyama,M.
TITLE        A novel human RasGAP-like gene that maps within the prostate cancer
             susceptibility locus at chromosome 1q25
JOURNAL      FEBS Lett. 441 (1), 127-131 (1998)
MEDLINE      99093006
REFERENCE    2  (bases 1 to 4368)
AUTHORS      Noto,S., Maeda,T., Hattori,S. and Hatakeyama,M.
TITLE        Direct Submission
JOURNAL      Submitted (11-FEB-1998) Department of Viral Oncology, The Cancer
             Institute, Japanese Foundation for Cancer Research, 1-37-1
             Kami-Ikebukuro, Toshima-ku, Tokyo, Japan
FEATURES             Location/Qualifiers
     source          1..4368
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1q24"
     CDS             126..3545
                     /note="ras GTPase activating protein"
                     /codon_start=1
                     /product="nGAP"
                     /protein_id="AAD04814.1"
                     /db_xref="GI:4105589"
                     /translation="MQTPEVPAERSPRRRSISGTSTSEKPNSMDTANTSPFKVPGFFS
                     KRLKGSIKRTKSQSKLDRNTSFRLPSLRSTDDRSRGLPKLKESRSHESLLSPCSTVEC
                     LDLGRGEPVSVKPLHSSILGQDFCFEVTYLSGSKCFSCNSASERDKWMENLRRTVQPN
```

```
ACCESSION   AF053938
VERSION     AF053938.1   GI:3213253
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 829)
  AUTHORS   Chen,H.-J. and Kennedy,M.B.
  TITLE     Identification and cloning of a novel 130 kd protein containing a
            ras GTPase-activating domain from the rat forebrain postsynaptic
            density
  JOURNAL   Abstr. - Soc. Neurosci. 23, 1466-1466 (1997)
REFERENCE   2  (bases 1 to 829)
  AUTHORS   Chen,H.J., Rojas-Soto,M., Oguni,A. and Kennedy,M.B.
  TITLE     A synaptic Ras-GTPase activating protein (p135 SynGAP) inhibited by
            CaM kinase II
  JOURNAL   Neuron 20 (5), 895-904 (1998)
  MEDLINE   98282016
REFERENCE   3  (bases 1 to 829)
  AUTHORS   Chen,H.-J. and Kennedy,M.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-1998) Biology, California Institute of
            Technology, 1200 E. California Blvd., MC 216-76, Pasadena, CA
            91125, USA
FEATURES             Location/Qualifiers
     source          1..829
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
     CDS             183..>829
                     /note="synaptic ras GTPase-activating protein; N-terminal
                     splice variant; synaptic ras-GAP"
                     /codon_start=1
                     /product="p135 SynGAP"
                     /protein_id="AAC23491.1"
                     /db_xref="GI:3213254"
                     /translation="MGLRPPTPTPSGGSGGSLPPPSHRQPLRRRCSSCCFPGEYHLG
                     RSRRKSVPGGKQYSMEAAPAAPFRPSQGFLSRRLKSSIKRTKSQPKLDRTSSFRQILP
                     RFRSADHDRARLMQSFKESHSHESLLSPSSAAEALELNDEDSIIKPVHSSILGQEFC
                     FEVTTSSGTKCFACRSRAAERDKWIENLQRAVKPNKDNSRRVDNVLKWIIEARELP"
BASE COUNT      182 a     266 c     218 g     163 t
ORIGIN

Query Match              11.1%;        Score 443.2;   DB 74;    Length 829;
Best Local Similarity    97.2%;        Pred. No. 3.7e-75;
Matches   451;  Conservative   0;  Mismatches  13;  Indels   0;  Gaps   0;
```

```
Qy    772  ggaagtggggggctctggggtatgggctcggggggagaggggggtcaggggggcggctca  831
Db    945  CCCAA-------------------------------------------------------  949

Qy    832  ggggggcaaagggaaaggaggctgtcctgctgtgcggctgaaggctgaaagccccgttaccagacaatg  891
           |||| ||||||| ||| | | ||||| |  ||| ||||||||||||  |||||| |||||||
Db    950  ----CAAAGGAAAGACAGAGGACCTTCATTCGGATTAAATCACGTTTCCAAACTATC  1004

Qy    892  agtatcctggccatggactatataaggagttgcagaatatgtgaccaaccactaccgc  951
           || |||||||||| ||||||| ||| ||| || ||  ||| | |||||||| ||||||
Db   1005  ACCATTCCTGCCTATGGACAATACAAGAGATTTGTCAGAATTGTCACCAGCAACTACACC  1064

Qy    952  atgctgtgtcgtgctggagcccgcccctcaatgtcaaggccaaggagagtcggctagt  1011
           |||| ||| ||| ||||| |||||||| ||| ||||  |||||| ||||| ||||| ||
Db   1065  ATGCTGTGTTCTGTTCCTTGAGCCAGTAATTAGTGTGAGAAATAAAGAGGAGTTGGCTTGT  1124

Qy   1012  gcactggttcacctcctgcaaagcacaggcaaggccaaggacttcctttcagacatggcc  1071
           || || |||| ||||| || ||  |||||| || || ||  |||||| |||| | |||||
Db   1125  GCCTTAGTGCACATTCTTCAAAGTACTGGCAGAGCAGCCAAGGATTTTCTGACTGACTTGGTG  1184

Qy   1072  atgtcagaggtagaccggttcatggagcgggaacacctcatattccgcgagaacacgctc  1131
           ||||| ||  ||| |||| ||| ||||| | ||| ||| ||||| |||||||||||| |||
Db   1185  ATGTCTGAGGTGGATCGTTGTGGAGGACATGAGTGTTGGAGCGAGAGGAGAACACTATT  1244

Qy   1132  gccactaagcatagaagagtatatgagactgattggccagaaatacctcaaggatgcc  1191
           |||| ||| || ||| |||||||||  | |||| |||||||||  | |||| |||| |||
Db   1245  GCCACCAAATCCATTGAGGAATACCTCAAGTTGTGGACAACAGTATCTTCATGACGCA  1304

Qy   1192  attggggagttcatccggctctgtatgaatctgaggagagaactgtgaagtagacccatc  1251
           || ||||| ||||| | | ||| |  ||||| |||| |||| || |||||| |||| ||
Db   1305  CTGGGGGAGTTTATCAAAGCTTGTGAGTGAGTGAGGAACTGTGAAGTGGATCCCAGC  1364

Qy   1252  aagtgcacagcgtccagtctgcagagcaccaggccaactgctcgtgatgtgctgctgttg  1311
           || ||| |||||| |||| | ||  |||| |||| | |  | | ||| || ||| ||||
Db   1365  AATGTTCATCTAGTGAACGATAGAGCCCATCAGAGCAACCTGAAAATGTGCTGTGAGCTG  1424

Qy   1312  gccctgtgcaaggtggtcaactccccattgcgtgttcccgaggaggctgaaggaggtgttt  1371
           || ||| | |||  | |||   | | ||||||| || ||| |||| || || |||| ||
Db   1425  GCTTTCTGCAAGATCAACATCAACTCTTACTGTTCCCTCGTCGTGAGTGAAAAGAAGTGTT  1484

Qy   1372  gcatcatggcggctgcgtgctgcaggagcggggcgggagcaagcaagcaatcagccgagagggctcatc  1431
           |||| ||||| || ||  | ||||||| |||| |||| |  | | | |||||| ||||| |||||
Db   1485  GCATCATGGAAGCAGCAGTGCCTGAACCGTGCGCAAGCAACGTGGACAATCAGCGAGAGGCTCATC  1544

Qy   1432  agcgcctgctcttcctcgcggctcctctgcggtcttggttgga  1491
           ||| |||| || | ||| || |||| | | ||| |  ||||| ||
Db   1545  AGTGCCTCATTATTTCTCCGTTTCTCCGTTTTCTGTGTCTCCCAGTGTCTTTCAAC  1604

Qy   1492  ctgatgcaggagataccagatgagcagacctcacgagaccctcaccctcatcgccaaggtt  1551
           || ||| ||||| | |  ||  |  | ||||| ||||  |||| |||| |||  ||||| |
Db   1605  CTTATGCAGGAGTATCCTGATGACCGCACATCTCGGACTCTAACTCTTATTGCCAAGGTC  1664

Qy   1552  atccagaacctgccaactttccaagttacctcaaaggagacttctctggggcttcatg  1611
           ||  |||| ||||||||| || ||||| |  ||||||||| || |||  || |||||
Db   1665  ATTCAGAACCTGCCAACTTTGCCAAGTTTGTAACAAGTTGGTAACAAAGAGGAATACATGGCATTCATG  1724

Qy   1612  aacgagtttctgagctggaatggggttcatcgcagcaattcttgtatgagatatccaac  1671
           |||| || ||  ||| ||   |||||| ||| | |||| |||  | ||| ||| | |||
Db   1725  AATGATTTTTAGAACATGAATGGGGTGGAATGAAGGCGCTTCTTCTTTTGGAGATCTTAAT  1784

Qy   1672  ctggacacactgaccaacgtgcggactcagtttgagggctacatagactggggcccgagctc  1731
           ||||| |||| | ||||| || || ||||  | |||| |  | |||||||||| ||||||
Db   1785  CCAGACACCATCTCAAACACCCCAGGCTTTGATGGTTACATTGATCTGGGCCCAGAGCTT  1844

Qy   1732  tccacacttcacgcccctgtctgaggtgctgcccagctcagcaagtg  1780
           |||| ||| |||  || || ||  || |||| || | ||| ||  |
Db   1845  TCAGTTTTGCATTCCTTACTGTGGGAAGTAGTTTCCCAACTTGATAAGG  1893

RESULT 10
AF053938
LOCUS       AF053938       829 bp       mRNA              ROD       12-JUN-1998
DEFINITION  Rattus norvegicus p135 SynGAP mRNA, partial cds.
```

```
Qy    385  aaatgqattqaqaattctacaqaqqqctqtgaaaccaacaaqqacaacaqccqccqqqta  444
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    726  AAATGGATTGAGAATCTACAGAGGGCTGTGAAACCCAACCAACAAGGACAACAGCCCGCGGGGTA  785

Qy    445  gataacgtgctgaaactatgatcatagaagctcgagagctgcc  488
                ||||||||||||||||||||||||||||||||||||||||||||
Db    786  GATAACGTGCTGAAACTATGGATCATAGAAGCTCGAGAGCTGCC  829
```

RESULT 11
```
AK024488
LOCUS       AK024488        4287 bp       mRNA        PRI        29-SEP-2000
DEFINITION  Homo sapiens mRNA for FLJ00087 protein, partial cds.
ACCESSION   AK024488
VERSION     AK024488.1  GI:10440489
KEYWORDS    fis (full insert sequence).
SOURCE      Homo sapiens adult spleen cDNA to mRNA, clone:as00087.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4287)
AUTHORS     Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE       The nucleotide sequence of a long cDNA clone isolated from human
            spleen
JOURNAL     Published Only in DataBase (2000) In press
REFERENCE   2  (bases 1 to 4287)
AUTHORS     Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE       Direct Submission
JOURNAL     Submitted (24-AUG-2000) to the DDBJ/EMBL/GenBank databases. Osamu
            Ohara, Kazusa DNA Research Institute, Department of Human Gene
            Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/NEDO,
            Tel:81-438-52-3913, Fax:81-438-52-3914)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing; Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing and clone selection:
            Kazusa DNA Research Institute.
FEATURES             Location/Qualifiers
     source          1..4287
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="as00087"
                     /dev_stage="adult"
                     /tissue_type="spleen"
                     /note="vector:pBluescriptII SK plus"
     gene            1339..3363
                     /gene="FLJ00087"
     CDS             <1339..3363
                     /gene="FLJ00087"
                     /note="For this clone, GeneMark analysis triggered an
                     alert for spurious CDS split. The result of GeneMark
                     analysis of the nucleotide sequence of this clone is
                     accessible through http://www.kazusa.or.jp/NEDO.
                     Start codon is not identified."
                     /codon_start=1
                     /product="FLJ00087 protein"
                     /protein_id="BAB15778.1"
                     /db_xref="GI:10440490"
                     /translation="GSTLHLRCLLWGANKPKRFPPSHSHLQDNVERERTWLSVWVHEA
                     KGLPRAAGAPGVRAELWLDGALLARTAPRAGPGQLFWAERFHFEALPPAARRLSLRLR
                     GLGPGSAVLGRVALAELELDAPRAPAAGLERWFPLLGAPAGAALRARIRARRLRVLPS
                     ERYKELAEFLTFHVARLCGALEPALPAQAKEELAAAMVRVLRATGRAQALVTDLGTAE
                     LARCGREALLFRENTLAYKAIDEYMKLVAQDYIQETLGQVVRRLCASTEDCEVDFSK
                     CPASELPEHQARLRNSCEEVFETIIHSYDWFPAELGIVFSSWREACKERGSEVLGPRL
                     VCASLFRLLCPAILAPSLFGLLAPDHPAPGPARTLTLIAKVIQNLANRAPFGEKRAYM
                     GFMNSFLEEHGPAMOCFLDQVAMVDVDAAPSGYOGSGDLALQLAVLHAQLCTIFAELD
                     QTTRDTLEPLLPTILRAIEIEGQPVLVSVPMRLPLPPAQVHSSLSAGEKPGFLAPRDLPK
                     HTPLISKSQSLRSVRRSESWARPRPDERPLRRPRPVQRTQSVPVRRPARRRQSAGPW
                     PRPKGSLSMGPRARPWTRDSASLPRKPSVPWQROMDQPQDRNQALGTHRPVNKLAE
                     LQCEVALREEQKVLSRLVESLSTQIRALTEQQELRGQLQDLDSRLRAG"

BASE COUNT       785 a      1529 c      1261 g      712 t
```

ORIGIN
```
Query Match             8.5%;  Score 336.4;  DB 34;  Length 4287;
Best Local Similarity  59.2%;  Pred. No. 7.7e-55;
Matches  574;  Conservative   0;  Mismatches 396;  Indels   0;  Gaps   0;

Qy    857  ctgctgtgcgctgaaggcccgctaccagacaatgagtatctgcccatggagctatata  916
                ||||||||| ||| ||| ||||| | |  |||||| || || |   |   ||||||||||
Db   1769  CAGCGGCTGCGGGCGCGGCGGATTCGGGCGCGTCGCTCGCCTGCGCGTGCCGTCCAGCGCTACA 1828

Qy    917  aggagtttgcagaatatgtgaccaaccactaccgcatgctgtgtgcgctggagcccg  976
                ||||| ||||| |||  | | | || |||   |   |||||| |  ||||||||||||
Db   1829  AGGAGCTGGCGGAGTTCCTCACCTTCCACTATGCGCGCCTGCTGTGCGGGGCCCTGGAGCCCG 1888

Qy    977  ccctcaatgtcaagggcaaggaggaggttcgctagtgcactgttcaatctgcaaagca 1036
                | |  |||||||| ||||||||||||| |||||||||||| |||| |||||||||||||
Db   1889  CGCTGCCTGCGCAGCCAAGGAGGAGTCGGCGCCAGGAGAGCCATGGTGCGCGTGCTGCGGGCCA 1948

Qy   1037  caggcaaggccaaggacttcctttcagacatgccatgtcaggataggaccggttcatgg 1096
                | || || ||| || |||||| | || |||| |||||||| || |||||||||||||||
Db   1949  CCGGCCCGGCGCCAGGCGCTGGTGACTGACCTGGCACTGCCAGCTGCCGCCCTGTTGGAG 2008

Qy   1097  agcggaacacctcatattccgcagaacacgctcgccactaaagccatagaagagtata 1156
                |  ||| || |||||| |  |||||| |  || ||||| |||||| |||| ||||| ||
Db   2009  GCCGTGAGGCGCTGCTGTCTGTTCCGGGAAAAACACATTGGCCACCAAGGCTATCGATGAGTACA 2068

Qy   1157  tgaactgattggccagaaatacctcaaggatgccattgggggagtcatccgggcctcgt 1216
                | || | ||  |||   || ||| |  || |||  | |||| || |||||||||||||||
Db   2069  TGAAGCTCGTGGCACAGGATTACCTCCAGGAGACCCTGGGACAGGTTGTGCGGCGTCT 2128

Qy   1217  atgaatctgaggagaactgtgaagtagaccccatcaagtgcacagcgtccagtctggcag 1276
                || |||  || |||||| ||||| |||  |||| || ||| ||| || |||||||||||
Db   2129  GTGCTTCTACTGAGGACTGTGAAGTGACCCCAGCAAATGTCCAGCCTCGGAGCTGCCAG 2188

Qy   1277  agcaccagccaacctgcggatgtgctgtgagttggcctgcaaggtgtcaactccc 1336
                || ||||||||| | || |||| ||||||||||   |||||||||  |   ||||||
Db   2189  AGCACCAGGCCAGACTTCCGGAACAGCTGCGGAGGAGGTCTTCGAAACCATTATCCATTCCT 2248

Qy   1337  attggtgttcccgaggggagctgaaggaggtgtttgcatcatggcgctgcgctgtcgag 1396
                || | |||| | |||| ||| ||||| |||    ||||| |  | ||| || || | ||
Db   2249  ACGACTGGTTCCCTGCGGAGCTGGGCATCGTGTTCTCAAGCTGGCGCGAGAAGCATGTAAAG 2308

Qy   1397  agcggggccgggaggacattgctgacaggctgatcagcgctcgcctcttcctgcgcttcc 1456
                | || |||| |||||  |||  |||||| ||||  || | || ||||||| || || ||||
Db   2309  AACGTGGCTCTGAGTGCGTGGGCCCCGACTGGGTGTGCGCCCTCTTCCGGGAGAACACCCTCTCC 2368

Qy   1457  tctgcccggccatcatgtcgccagtcgtgttggactgatgcaggagtaccagagtgagc 1516
                | | |||||||| ||| | |||||| | |||||||||| ||||||||| |||||||||||
Db   2369  TGTGCCTGCCATCCTGGCCAGCCTCTTGTGGTTTGGTTTGGCACCAGACCATCCAGCACACCCG 2428

Qy   1517  agacctcacgaacccaccctcatcgccaaggttatccaagacctggccaactttcca 1576
                || || |||||||| |||||||| || ||| | ||| ||||| |||| || ||||| | ||
Db   2429  GCCCAGCCCGCACCCTCCACATTGCCATCCCAGAGTCATCCAGAACCTCCGCAACCGTGCCC 2488

Qy   1577  agtttacctcaaagggagactcctggccttcatgaacgagttctggagctggaatggg 1636
                ||||| |||  ||||||| | || |||| |||||||  | |||||| |||||||||||||
Db   2489  CGTTCCGTGAGAAGGAGGACGGCCTACATGGGCTTCATGAATAGCTTCCTGGAGAACATGGAC 2548

Qy   1637  gttctatgcgacaattcttgtatgagatatccaacctggacaacactgaccaacagcaca 1696
                |  ||| || |||  | | || | |||  ||  | ||  |||| |||| || || ||||||
Db   2549  CAGCCATGCAATGCTTCCTGGACCAGGTAGCCATGGTGGATGTGGATGCCGCCCCCAGTG 2608

Qy   1697  gtttgagggctacatagactggggccgcgagctctcacacttcacgccctgtctgqq  1756
                ||||| || |||  | ||| | |||||||| || ||| ||||| | ||| || |||||
Db   2609  GTTACCAGGGCAGTGGAGATGGGCAGTGGTGATCTGGCGCTTCAGCTGTCCGTCCCAGCTCGTA 2668

Qy   1757  aggtgctgcccagctcagcagaagcctctgaagctggcccgctgcccccggctcc  1816
                ||  ||||||||| |||| || |||    || || || ||| |   |  || |||||
Db   2669  CAATTTTGCTGAGCTTGACCAGAACACCCGAGACACCCCCGGAACCACTGCCCACCATCC 2728

Qy   1817  tcagcgacat 1826
                | | ||||
```

Db   2729 TGCGCAGCCAT 2738

RESULT 12
AC019800/c
LOCUS   AC019800   56366 bp   DNA   HTG   03-JAN-2000
DEFINITION   Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
             pieces.
ACCESSION   AC019800
VERSION   AC019800.1 GI:6665097
KEYWORDS   HTG; HTGS_PHASE2.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
             Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
             Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
             Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 56366)
AUTHORS   Adams,M. and Venter,J.C.
TITLE   Direct Submission
JOURNAL   Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
             Rockville, MD, USA
COMMENT   This sequence was identified as CDM:10210991 by the submitter.
             For more information on this record e-mail to fly@celera.com.
             * NOTE: This is a 'working draft' sequence.
             * This sequence will be replaced
             * by the finished sequence as soon as it is available and
             * the accession number will be preserved.
FEATURES   Location/Qualifiers
source   1..56366
             /organism="Drosophila melanogaster"
             /db_xref="taxon:7227"
BASE COUNT   16234 a   11924 c   11808 g   16400 t
ORIGIN

Query Match   6.0%;   Score 240.2; DB 39; Length 56366;
Best Local Similarity 50.6%;   Pred. No. 1.2e-36;
Matches   748; Conservative   0; Mismatches 683; Indels 48; Gaps 5;

Qy   376 gaaaggacaaatggattgagaatctacagagggctgtgaaaccacaaggacaacagc 435
           || |||| || || ||| | | | | |||||||||| ||||||
Db 25786 GAGCGGCGACCTTTGGATCTACTCGGTGCCAAGTCGAATCGTCTCGGAATGCAGAGCACACG 25727

Qy   436 cgccgggtagataacgtgctgaaactatggatcatatagaagtcgagagagtgccccccag 495
           || ||||| | ||| | ||| || | ||||| ||| | | |||||
Db 25726 CGTCGCACGGACAACTCCGTGAAGATGTGGGTGTACGAGGCGAAAAATCTGCCGCCCAAG 25667

Qy   496 aagcgatattactgcgagttatgcctggacgacatgctcttatgcacggaccacttccag 555
           |||||||| | ||||| |||||||| ||||| |||| ||
Db 25666 AAGCGTTACTTTTGCGAACTGCAATTGGACAAGACCGCTGTACGGCCGGACTTCGGTGAAG 25607

Qy   556 cccggctcagcctcaggagacactgtcttttgggcgcagcacttcgagtttaacaaactg 615
           |||||||| ||| ||| | || | ||| | | | ||||||
Db 25606 CTGCAGAC-------GGATCTGCTGTTTTGGGGGAGCACTTCGATTCCCCGACATA 25556

Qy   616 cctgctgtccgggcgcgggctgcatctgcatctgactcgtgaccggacaaaaagcggaagaag 675
           || ||| ||| || |||| | || | ||||| || ||| |||| | |||| |
Db 25555 CCCAGATTAATGCATCACTGTTAACGTTTTCCGTGAGGTGGACAAGAAGAAGAAGCCG 25496

Qy   676 gacaaggcaggctacgttggcctggtgactgttccagtggccaccctggcgcgccac 735
           ||||| || ||||| |||| | |||| | ||| |||| ||| |||| ||
Db 25495 GACAAATACCAATTTGTGGGATCGGTCGAAGATACCCCGTGCACACGATGTCACCTCCAGATTG 25436

Qy   736 ttcacagagcagtggtacccctgaccctgccaacaggaagtggggggctctgggggtatg 795
           ||||||| || || || || || | ||||| ||||||| | || | || |||
Db 25435 CCCTTGCGAACAATGGTATCCCATA-------------------CTCGAGCGACAAGGCGGGC 25394

Qy   796 ggctcggggggagggagggggtcagggggcggctcagggggcaaggggaagggaggctgt 855
           || || ||| || ||||||| ||||| | || |||||| |||||| | | ||
Db 25393 GACAGTCTGGGCAGGACATCGGGCGGCGGCGGCCAGTGGGTCCAAGGAGGTCCAAGGACAAGGAGCAATTG 25334

Qy   856 cctgctgtgcggctgaaggccgttaccagacaatgagtatcctgccatggagctatat 915
           || || || || || | | | ||| || |||||||| | | |
           |||||||||||||| ||

Db   2729 [continued]

Db 25333 CCCACGCTGGAGGATCAAGTGTCGTTCCAGAGCACCGACATCCTGCCCATCAATGTGTAC 25274

Qy   916 aaggagtttgcagaatatgtgaccaaccactaccgcatgctgtgtgccgtggagccc 975
           || ||| || || ||| |||| || ||| ||
Db 25273 GGCAACTTTTGACGTACCTCAAGGAGAACTATAAGCCGTGTGCGAGACCCTGGAGCCG 25214

Qy   976 gccccaatgtcaaggcaagggaggagtcgctagtccgactggttcacatcctgcaaagc 1035
           ||||| || | ||| || ||| || | || || ||
Db 25213 GTGATCGGAGTCAAGGCCAAGGAGGACATTGGACAGGCACTGGCTGCTGCTGATGCACGCA 25154

Qy   1036 acagcaaggcaaggacttcctttcagacatggccatgtcagatgcagaggtagagaccggttcatg 1095
           | | || | | | |||| || ||
Db 25153 CAGGGATTGGCGGGCGCCTTCCTCCACCGATGTGTGGCCCTCGATCGTGCTGCTGGAGTTGGC 25094

Qy   1096 gagcgggaacaccctcatattccgcgagaacacgctcgccactaaagccatagaagagtat 1155
           | | ||| || | | ||||| || | || | | || |
Db 25093 GATCAGAGG---CTTACGTTCAGGGCAACTCCTTGGCCACCAAGACATGAGGGCATTC 25037

Qy   1156 atgagactgattggccagaaatacctcaaggatgccattgggagttcatccgggctcty 1215
           | | |||| | || || || | |||| |||| ||
Db 25036 CTCAAGCTGACGGGCGAACAGTATCTGCAGGACACACTATCCGCACCAATAAACGAGCTA 24977

Qy   1216 tatgaatctgaggagaactgtgaagtagacccatcaagtca---------cagcgtcc 1266
           | | |||| || | || || || | | |||| |||
Db 24976 ATTCAGTCGGAGAGAGGGACTGCGAGGTGGATCCCCACCAAGACGAGCCGGTTCGTCGGCCGGGT 24917

Qy   1267 agtctggcagacaccagccaacctgcggatgtgctgtgagtcgtgagtggcctgtcgcaagtg 1326
           || |||||| | | | || || | | | | ||| || ||
Db 24916 TCGCTGCAGCGACGACAGCAGGCGCCTTGCGTGCGGCGGTCGGGGGCGGGCGTGGCAGTGCATC 24857

Qy   1327 gtcaactcccattgctgttccgggcggtgaaggagggtgtttgcatcatggcggctg 1386
           || | | || || | | ||| || ||| | | | |
Db 24856 TTCGAATCGCACAAGCATTTTCCCGCCCAGTTACGAAAATTGCTTTGGCGACGTTCCGGGAG 24797

Qy   1387 cgctgtgcagacgggcgggacattggcgggaggacaattc tgacaggctgatcagcgcctcgcctctc 1446
           | | | | | ||||| | | || || |||
Db 24796 CGCTTGCAGCAGCTGGGCCCTCAGGATATGGCCGCACAACCTGATCTCCGGCGAGCATTTTTC 24737

Qy   1447 ctgcgcttcctgcgctggcgtcatcgtgcgccacgtcgttggactgatgcaggaggtac 1506
           |||||||| |||||| || || || | ||
Db 24736 CTGCGCTTCCTGTGCCGGCGACCATCCGTCGCCGTCTTCAATATCACCAGCGAACTG 24677

Qy   1507 ccagatgagcagacctcacgaacctcacccctcatcgccaaggttatccagaaactggcc 1566
           || || | ||||| | | ||||| | || || || ||| | |||
Db 24676 CCGTCCGCACGGCCACCCGCAATCTCACACTGTGGCCAAGACCCTGCCAAACATTGGCC 24617

Qy   1567 aacttttccaagttacctcaaaggaggacttcctggcttcatgaacgagtttctgggag 1626
           || || || || | || || ||| ||| | |||| || |||||
Db 24616 AATTCACCCGCCTTCCAGGCAAGAGAACTTTATGGAGTTTCTCAACGATTTCTTCGAG 24557

Qy   1627 ctggaatggggttctatgcagcaattcttgtatgagatatccaacctggacacactgac- 1685
           || ||| ||| ||| |||| | |||| || | ||| | ||
Db 24556 CAGGAGGCCGCTGCGGAGCGCAACAGTTTCTGGAGATTATATCCACCGCCGAGCACACCCA 24497

Qy   1686 ------caacagcagtttgaggctacatagactggccgcgagctctccaca 1737
                 | |||| | | | | |||| || || ||
Db 24496 GCTCCAGACTCGATCCTCGATTGGGCCGGGTACATCGACCAGGGCAAACAGTTGTCCATA 24437

Qy   1738 cttcacgccctgctctgggagggtgctgcccagctcagcaaggaagccctcctgaagctg 1797
           || ||| || | | |||| || | ||
Db 24436 CTACACAGTTGCTCAGCGAAAGCCTGGCCAAGCTGCCGCGGAGCGCCAGCGACGACGAGCTG 24377

Qy   1798 ggcccgctgccccggctcctccagccgacatcagcacagcc 1836
           ||| || | || | ||| || ||
Db 24376 GATCCGTTGCGACGCAATATTCCGATGAAATCAGCCGAGCC 24338

RESULT 13
AC012162
LOCUS   AC012162   191504 bp   DNA   HTG   09-FEB-2000
DEFINITION   Drosophila melanogaster chromosome x clone BACR01N10 (D1115)
             RPCI-98 01.N.10 map 16D-16F strain y; cn bw sp, *** SEQUENCING IN
             PROGRESS ***, 142 unordered pieces.

```
ACCESSION   AC012162
VERSION     AC012162.9  GI:6957580
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 191504)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R. Karra,K.,
            Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
            Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
            Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D.,
            Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 191504)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R. Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Feb 11, 2000 this sequence version replaced gi:6833825.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
            the following cutoffs: length >= 200 bases.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 142 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    456: contig of 456 bp in length        *     8506    8585: gap of unknown length
            *      457    536: gap of unknown length             *     8586    9227: contig of 642 bp in length
            *      537   1004: contig of 468 bp in length        *     9228    9307: gap of unknown length
            *     1005   1084: gap of unknown length             *     9308   10091: contig of 784 bp in length
            *     1085   1958: contig of 874 bp in length        *    10092   10171: gap of unknown length
            *     1959   2038: gap of unknown length             *    10172   11099: contig of 928 bp in length
            *     2039   2248: contig of 210 bp in length        *    11100   11179: gap of unknown length
            *     2249   2328: gap of unknown length             *    11180   11747: contig of 568 bp in length
            *     2329   2858: contig of 530 bp in length        *    11748   11827: gap of unknown length
            *     2859   2938: gap of unknown length             *    11828   12767: contig of 940 bp in length
            *     2939   3325: contig of 387 bp in length        *    12768   12847: gap of unknown length
            *     3326   3405: gap of unknown length             *    12848   13398: contig of 551 bp in length
            *     3406   4165: contig of 760 bp in length        *    13399   13478: gap of unknown length
            *     4166   4245: gap of unknown length             *    13479   14007: contig of 529 bp in length
            *     4246   4716: contig of 471 bp in length        *    14008   14087: gap of unknown length
            *     4717   4796: gap of unknown length             *    14088   14867: contig of 780 bp in length
            *     4797   5373: contig of 577 bp in length        *    14868   14947: gap of unknown length
            *     5374   5453: gap of unknown length             *    14948   15562: contig of 615 bp in length
            *     5454   6050: contig of 597 bp in length        *    15563   15642: gap of unknown length
            *     6051   6130: gap of unknown length             *    15643   16859: contig of 1217 bp in length
            *     6131   6720: contig of 590 bp in length        *    16860   16939: gap of unknown length
            *     6721   6800: gap of unknown length             *    16940   17986: contig of 1047 bp in length
            *     6801   7461: contig of 661 bp in length        *    17987   18066: gap of unknown length
            *     7462   7541: gap of unknown length             *    18067   19341: contig of 1275 bp in length
            *     7542   8104: contig of 563 bp in length        *    19342   19421: gap of unknown length
            *     8105   8184: gap of unknown length             *    19422   20926: contig of 1505 bp in length
            *     8185   8505: contig of 321 bp in length        *    20927   21006: gap of unknown length
                                                                 *    21007   22153: contig of 1147 bp in length
                                                                 *    22154   22233: gap of unknown length
                                                                 *    22234   23236: contig of 1003 bp in length
                                                                 *    23237   23316: gap of unknown length
                                                                 *    23317   24607: contig of 1291 bp in length
                                                                 *    24608   24687: gap of unknown length
                                                                 *    24688   25620: contig of 933 bp in length
                                                                 *    25621   25700: gap of unknown length
                                                                 *    25701   26300: contig of 600 bp in length
                                                                 *    26301   26380: gap of unknown length
                                                                 *    26381   27488: contig of 1108 bp in length
                                                                 *    27489   27568: gap of unknown length
                                                                 *    27569   28215: contig of 647 bp in length
                                                                 *    28216   28295: gap of unknown length
                                                                 *    28296   29440: contig of 1145 bp in length
                                                                 *    29441   29520: gap of unknown length
                                                                 *    29521   30587: contig of 1067 bp in length
                                                                 *    30588   30667: gap of unknown length
                                                                 *    30668   32042: contig of 1375 bp in length
                                                                 *    32043   32122: gap of unknown length
                                                                 *    32123   33740: contig of 1618 bp in length
                                                                 *    33741   33820: gap of unknown length
                                                                 *    33821   35375: contig of 1555 bp in length
                                                                 *    35376   35455: gap of unknown length
                                                                 *    35456   36258: contig of 803 bp in length
                                                                 *    36259   36338: gap of unknown length
                                                                 *    36339   38003: contig of 1665 bp in length
                                                                 *    38004   38083: gap of unknown length
                                                                 *    38084   39509: contig of 1426 bp in length
                                                                 *    39510   39589: gap of unknown length
                                                                 *    39590   40719: contig of 1130 bp in length
                                                                 *    40720   40799: gap of unknown length
                                                                 *    40800   42453: contig of 1654 bp in length
                                                                 *    42454   42533: gap of unknown length
                                                                 *    42534   43841: contig of 1308 bp in length
                                                                 *    43842   43921: gap of unknown length
                                                                 *    43922   45410: contig of 1489 bp in length
                                                                 *    45411   45490: gap of unknown length
                                                                 *    45491   46782: contig of 1292 bp in length
                                                                 *    46783   46862: gap of unknown length
                                                                 *    46863   48570: contig of 1708 bp in length
                                                                 *    48571   48650: gap of unknown length
                                                                 *    48651   51042: contig of 2392 bp in length
                                                                 *    51043   51122: gap of unknown length
                                                                 *    51123   52760: contig of 1638 bp in length
                                                                 *    52761   52840: gap of unknown length
```

```
    *     52841     54712: contig of 1872 bp in length
    *     54713     54792: gap of unknown length
    *     54793     56521: contig of 1729 bp in length
    *     56522     56601: gap of unknown length
    *     56602     58938: contig of 2337 bp in length
    *     58939     59018: gap of unknown length
    *     59019     60837: contig of 1819 bp in length
    *     60838     60917: gap of unknown length
    *     60918     62988: contig of 2071 bp in length
    *     62989     63068: gap of unknown length
    *     63069     65583: contig of 2515 bp in length
    *     65584     65663: gap of unknown length
    *     65664     68774: contig of 3111 bp in length
    *     68775     68854: gap of unknown length
    *     68855     71235: contig of 2381 bp in length
    *     71236     71315: gap of unknown length
    *     71316     74113: contig of 2798 bp in length
    *     74114     74193: gap of unknown length
    *     74194     77405: contig of 3212 bp in length
    *     77406     77485: gap of unknown length
    *     77486     80936: contig of 3451 bp in length
    *     80937     81016: gap of unknown length
    *     81017     83357: contig of 2341 bp in length
    *     83358     83437: gap of unknown length
    *     83438     87456: contig of 4019 bp in length
    *     87457     87536: gap of unknown length
    *     87537     91532: contig of 3996 bp in length
    *     91533     91612: gap of unknown length
    *     91613     97143: contig of 5531 bp in length
    *     97144     97223: gap of unknown length
    *     97224    105593: contig of 8370 bp in length
    *    105594    105673: gap of unknown length
    *    105674    112071: contig of 6398 bp in length
    *    112072    112151: gap of unknown length
    *    112152    121011: contig of 8860 bp in length
    *    121012    121091: gap of unknown length
    *    121092    143113: contig of 22022 bp in length
    *    143114    143193: gap of unknown length
    *    143194    143587: contig of 394 bp in length
    *    143588    143667: gap of unknown length
    *    143668    143927: contig of 260 bp in length
    *    143928    144007: gap of unknown length
    *    144008    144842: contig of 835 bp in length
    *    144843    144922: gap of unknown length
    *    144923    145360: contig of 438 bp in length
    *    145361    145440: gap of unknown length
    *    145441    145915: contig of 475 bp in length
    *    145916    145995: gap of unknown length
    *    145996    146532: contig of 537 bp in length
    *    146533    146612: gap of unknown length

  Query Match             6.0%;  Score 240.2;  DB 36;  Length 191504;
  Best Local Similarity   50.6%;  Pred. No. 8.4e-37;
  Matches  748;  Conservative    0;  Mismatches  683;  Indels   48;  Gaps    5;

Qy    376 gaaagggacaaatggattgagaatctacagagggctgtgaaacccaacaaggacaacagc 435
          ||  | |||   |||||  |    || |  | | |  |    || ||    ||  |||
Db 113379 GAGCGCGACCTTTGGATCTACTCGCTGCGCAAGTCGATCGCTCCGAATGCAGAGCACACG 113438

Qy    436 cgccgggtagataacgtgctgaaactatggatcatagaagctcgagagctgccccccaag 495
          || ||    || |||  ||||||  | ||| |    || ||   | | ||||| ||||||
Db 113439 CGTCGCACGGACAACTCGCTGAAGATGTGGGTGTACGAGGCGAAAAATCTGCCGCCCAAG 113498

Qy    496 aagcgatattactgcgagttatgcctggacgacatgctctatgcacggaccacttccaag 555
          ||||| || |  |||||  |     ||||| | | ||| || |  |||||  |    |||
Db 113499 AAGCGTTACTTTTGCGAACTGCAATTGGACAAGACGCTGTACGGCCGGACTTCGGTGAAG 113558

Qy    556 ccccgctcagcctcaggagacactgtcttttggggcgagcacttcgagtttaacaacctg 615
          |  |   |          ||    | ||||||||| ||||||||||| ||   | || |
Db 113559 CTGCAGAC---------GGATCTGCTGTTTTGGGGGGAGCACTTCGATTTCCCCGACATA 113609

Qy    616 cctgctgtccgggcgctgcggctgcatctgtaccgtgactcggacaaaaagcggaagaag 675
          || |   |    | | |     |  |  | | ||||||   |||||| |||  ||||  |
Db 113610 CCCGAGATTAATGTGATCACTGTTAACGTTTTCCGTGAGGTGGACAAGAAGAAGAAGCGG 113669

Qy    676 gacaaggcaggctacgttggcctggtgactgttccagtggccaccctggctgggcgccac 735
          |||||       |  || ||   |||||   | || |||  |    |  |     |
Db 113670 GACAAATACCAATTTGTGGGATCGGTGAAGATACCCGTGCACGATGTCACCTCCAGATTG 113729

Qy    736 ttcacagagcagtggtaccccgtgaccctgccaacaggaagtgggggctctgggggtatg 795
            |   || || ||||| ||| |                     | ||     ||
Db 113730 CCCTGCGAACAATGGTATCCCATA------------------CTGAGCGACAAGGCGGGC 113771

Qy    796 ggctcggggggaggaggggggtcagggggcggctcagggggcaaagggaaaggaggctgt 855
          | |     |||  |      |   || ||||||   |||  ||| |  || |     |
Db 113772 GACAGTCTGGGCAGGACATCGGGCGGCGGCGGCAGTGGGTCCAAGGACAAGGAGCAATTG 113831

Qy    856 cctgctgtgcggctgaaggcccgttaccagacaatgagtatcctgcccatggagctatat 915
          ||  |  || || | |||   |||| |||||  |     |||||||||||  |  | ||
Db 113832 CCCACGCTGAGGATCAAGTGTCGTTTCCAGAGCACCGACATCCTGCCCATCAATGTGTAC 113891

Qy    916 aaggagtttgcagaatatgtgaccaaccactaccgcatgctgtgtgccgtgctggagccc 975
              | |||      ||  | |   |  ||||        |||| |     ||||||||
Db 113892 GGCAACTTTTTGACGTACCTCAAGGAGAACTATAAGCGCGTGTGCGAGACCCTGGAGCCG 113951

Qy    976 gccctcaatgtcaagggcaaggaggaggtcgctagtgcactggttcacatcctgcaaagc 1035
          |   ||   ||||||| |||||||||  | |     |||||||| |   |  ||||
Db 113952 GTGATCGGAGTCAAGGCCAAGGAGGACATTGGACAGGCACTGGTGCTGCTGATGCACGCA 114011

Qy   1036 acaggcaaggccaaggacttcctttcagacatggccatgtcagaggtagaccggttcatg 1095
             ||   |||    | ||||||  | ||  |||        ||  |     ||  |
Db 114012 CAGGGATTGGCGGGCGCCTTCCTCACCGATGTGGTGGCCCTCGATCTGCTGCGAGTTGGC 114071

Qy   1096 gagcgggaacacctcatattccgcgagaacacgctcgccactaaagccatagaagagtat 1155
          || | |      || |  ||| | |  ||| |  | ||||| ||   ||| || |  |
Db 114072 GATCAGAGG---CTTACGTTCAGGGGCAACTCCTTGGCCACCAAGAGCATGGAGGCATTC 114128

Qy   1156 atgagactgattggccagaaatacctcaaggatgccattggggagttcatccgggctctg 1215
           | |  ||||  ||| |  | || ||  ||||  |  |    |     ||    |  ||
Db 114129 CTCAAGCTGACGGGCGAACAGTATCTGCAGGACACACTATCCGCACCAATAAACGAGCTA 114188

Qy   1216 tatgaatctgaggagaactgtgaagtagaccccatcaagtgca---------cagcgtcc 1266
            | | || |||  | |||| || || || |||| ||||   |         | |||
Db 114189 ATTCAGTCGGAGAGGGACTGCGAGGTGGATCCCACCAAGACGAGCGGTTCGTCGGCGGGT 114248

Qy   1267 agtctggcagagcaccaggccaacctgcggatgtgctgtgagttggccctgtgcaaggtg 1326
             |||      || ||||||  | ||||               |||   |       |
Db 114249 TCGCTGCAGCGACAGCAGGCCGCCTTGCGTGGCGCGGTCCGAGGGGCGTGGCAGTGCATC 114308

Qy   1327 gtcaactcccattgcgtgttcccgagggagctgaaggaggtgtttgcatcatggcggctg 1386
           || | || ||       || ||     || |     |    |||||  | |  |||  |
Db 114309 TTCGAATCGCACAAGCATTTTCCCGCCCAGTTACGAAATTGCTTTGCGACGTTCCGGGAG 114368

Qy   1387 cgctgtgcagagcggggccgggaggacattgctgacaggctgatcagcgcctcgctcttc 1446
          ||||      ||| ||||||  |||| || || ||||  ||||||   ||     | |||
Db 114369 CGCTTGCAGCAGCTGGGCCGTCAGGATATGGCCGACAACCTGATCTCGGCGAGCATTTTC 114428

Qy   1447 ctgcgcttcctctgcccggccatcatgtcgcccagtctgtttggactgatgcaggagtac 1506
          ||||||||||| |||||||||||| |||||||    || ||     | |     ||
Db 114429 CTGCGCTTCCTGTGCCCGGCCATCCTGTCGCCGTCGCTCTTCAATATCACCAGCGAACTG 114488

Qy   1507 ccagatgagcagacctcacgaaccctcaccctcatcgccaaggttatccagaacctggcc 1566
          ||    |  | | || | || |  ||||| ||  | ||||||    | || |   |||||
Db 114489 CCGTCCGCACGGGCCACCCGCAATCTCACACTGGTGGCCAAGACCCTGCAAACATTGGCC 114548

Qy   1567 aacttttccaagtttacctcaaaggaggacttcctgggcttcatgaacgagtttctggag 1626
          || ||  ||   ||       || ||| ||||  |||  ||  | ||||| || || |||
Db 114549 AATTTCACCCGCTTCCAGGGCAAAGAGAACTTTATGGAGTTTCTCAACGATTTCCTCGAG 114608

Qy   1627 ctggaatggggttctatgcagcaattcttgtatgagatatccaacctggacacactgac- 1685
          | |||    | |   ||||| || ||  || |    |||||||  | |      |   |
Db 114609 CAGGAGGCCGCTCGCATGCAACAGTTTCTGGAGATTATATCCACGCGGCCGGAGCACCCA 114668

Qy   1686 --------caacagcagcagttttgagggctacatagacttgggccgcgagctctccaca 1737
                  |   | |  |  ||  |  || ||||| |||  ||||    || | |||| |
Db 114669 GCTCCAGACTCGATCCTCGATTGGGCCGGGTACATCGACCAGGGCAAACAGTTGTCCATA 114728
```

```
Qy   1738  cttcacgcccgtctctgggagtgtgtgcccagctcagcaagaggaagcctcctgaagctg  1797
                 ||||  ||||   |||||  || |||||   ||||| |||     ||||    ||||
Db 114729  CTACACAGTTTGCTCAGCGAAAGCCTGGCCAAGCTGGCCGAGGCCAGGCCAGCACGAGCTG 114788

Qy   1798  ggcccgctgcccggcctcctcagcgacatcagcacagcc 1836
                 | |||| |||| ||  | ||||||  || |||||| ||||
Db 114789  GATCCGTTGCAGCATATTCTCGATGAAATCAGCCGAGCC 114827
```

RESULT 14
ACO12161/c
LOCUS   ACO12161   202741 bp   DNA   HTG   16-NOV-1999
DEFINITION Drosophila melanogaster chromosome X clone BACR01I22 (D1114)
            RPCI-98 01.I.22 map 16B-16D strain y; cn bw sp; *** SEQUENCING IN
            PROGRESS ***, 142 unordered pieces.
ACCESSION   ACO12161
VERSION   ACO12161.7 GI:6435885
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 202741)
AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K.,
            Kearney,L., Lee.B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R.,
            Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B.,
            Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D.,
            Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.
TITLE   Sequencing of Drosophila melanogaster
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 202741)
AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee.B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,H., Zieran,L.L. and
            Rubin,G.M.
TITLE   Direct Submission
JOURNAL   Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT   On Nov 16, 1999 this sequence version replaced gi:6425635.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu. All contigs in this submission meet
            the following cutoffs: length >= 200 bases.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 142 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *   1   509: contig of 509 bp in length
            *   510   589: gap of unknown length
            *   590 1201: contig of 612 bp in length
            *   1202 1281: gap of unknown length
            *   1282 2394: contig of 1113 bp in length
            *   2395 2474: gap of unknown length
            *   2475 3512: contig of 1038 bp in length
            *   3513 3592: gap of unknown length
            *   3593 4500: contig of 908 bp in length
            *   4501 4580: gap of unknown length
            *   4581 5465: contig of 885 bp in length
            *   5466 5545: gap of unknown length
            *   5546 6253: contig of 708 bp in length
            *   6254 6333: gap of unknown length
            *   6334 7105: contig of 772 bp in length
            *   7106 7185: gap of unknown length
            *   7186 8077: contig of 892 bp in length
            *   8078 8157: gap of unknown length
            *   8158 9047: contig of 890 bp in length
            *   9048 9127: gap of unknown length
            *   9128 9530: contig of 403 bp in length
            *   9531 9610: gap of unknown length
            *   9611 10173: contig of 563 bp in length
            *   10174 10253: gap of unknown length
            *   10254 10841: contig of 588 bp in length
            *   10842 10921: gap of unknown length
            *   10922 11807: contig of 886 bp in length
            *   11808 11887: gap of unknown length
            *   11888 13165: contig of 1278 bp in length
            *   13166 13245: gap of unknown length
            *   13246 14526: contig of 1281 bp in length
            *   14527 14606: gap of unknown length
            *   14607 15801: contig of 1195 bp in length
            *   15802 15881: gap of unknown length
            *   15882 16640: contig of 759 bp in length
            *   16641 16720: gap of unknown length
            *   16721 17679: contig of 959 bp in length
            *   17680 17759: gap of unknown length
            *   17760 18797: contig of 1038 bp in length
            *   18798 18877: gap of unknown length
            *   18878 19955: contig of 1078 bp in length
            *   19956 20035: gap of unknown length
            *   20036 20931: contig of 896 bp in length
            *   20932 21011: gap of unknown length
            *   21012 21643: contig of 632 bp in length
            *   21644 21723: gap of unknown length
            *   21724 23416: contig of 1693 bp in length
            *   23417 23496: gap of unknown length
            *   23497 24457: contig of 961 bp in length
            *   24458 24537: gap of unknown length
            *   24538 25955: contig of 1418 bp in length
            *   25956 26035: gap of unknown length
            *   26036 27838: contig of 1803 bp in length
            *   27839 27918: gap of unknown length
            *   27919 29686: contig of 1768 bp in length
            *   29687 29766: gap of unknown length
            *   29767 30976: contig of 1210 bp in length
            *   30977 31056: gap of unknown length
            *   31057 32460: contig of 1404 bp in length
            *   32461 32540: gap of unknown length
            *   32541 34644: contig of 2104 bp in length
            *   34645 34724: gap of unknown length
            *   34725 36341: contig of 1617 bp in length
            *   36342 36421: gap of unknown length
            *   36422 37864: contig of 1443 bp in length
            *   37865 37944: gap of unknown length
            *   37945 39424: contig of 1480 bp in length
            *   39425 39504: gap of unknown length
            *   39505 41341: contig of 1837 bp in length
            *   41342 41421: gap of unknown length
            *   41422 42970: contig of 1549 bp in length
            *   42971 43050: gap of unknown length
            *   43051 44419: contig of 1369 bp in length
            *   44420 44499: gap of unknown length
            *   44500 46484: contig of 1985 bp in length
            *   46485 46564: gap of unknown length
            *   46565 48446: contig of 1882 bp in length
            *   48447 48526: gap of unknown length
            *   48527 50693: contig of 2167 bp in length
            *   50694 50773: gap of unknown length
            *   50774 53578: contig of 2805 bp in length
            *   53579 53658: gap of unknown length
            *   53659 55753: contig of 2095 bp in length
            *   55754 55833: gap of unknown length
```

```
Db 126449  GAGCCGACCTTTGGATCTACTCGCTGCGCAAGTCGATCGCTCCGAATGCAGAGCACACG  126390
Qy    436  cgccgggtagataacgtgctgaaactatggatcatagaagtcgagagctgccccccaag  495
           ||| |   |||||||| ||||||| || | |||||||  | ||||| ||| |||||||
Db 126389  CGTCCACGGACAACTCGCTGAGAGATGTGGGTGTACGAGGCGAAAATCTGCCGCCCAAG  126330
Qy    496  aagcgatattactgcgagttatgccgacgacatgctcatgcacggaccacttccaag  555
           ||||| || || |||||| || ||| |||||| |||| | ||| || |||||||
Db 126329  AAGCCTTACTTTTGCGAACTGCAATTGGACAAGAGCGCTGTACGGCCGGACTTCGGTGAAG  126270
Qy    556  ccccgctcagcctcaggagacactgtctttgggcgagcacttcgagttaacaactg  615
           ||  |||||| || |||| |  || ||||||||||||||||||| || ||||||||
Db 126269  CTGCAGAC--------GGATCTGCGTGTTTTGGGGGAGCACTTCGATTTCCCCGACATA  126219
Qy    616  cctgctgtccgggcgctcgggctgcatctgacctgtactgacaaaaagcggaagag  675
           |  || ||||||||| ||||| |||| || ||||||||||||| |||||||||| |
Db 126218  CCCGAGATTAATGTGATCACTGTTAACGTTTCCGTGAGGTGGACAAGAAGAAGAAGCGG  126159
Qy    676  gacaaggcaggctacgttggcctggtgactgttccagtggccacctggctgggcgccac  735
           |||||  |  || || ||| | ||||||||||||||| ||||||||||||||||| ||
Db 126158  GACAAATACCAATTTGTGGGATCGGTGAAGATACCCGTGCACGATGTCACCTCCAGATTG  126099
Qy    736  ttcacagacgagtggtaccccgtgaccctgcaacagagaagtggggctctgggggtaty  795
           ||  | || || | |||  ||||| |||| ||| || ||||||||| |||
Db 126098  CCCTGCGAACAATGGTATCCCATA----------CTGAGCGCAAGGCGGGC  126057
Qy    796  ggctcggggggaggagggggtcaggggcggccaggggcaaagggaaagaggagctgt  855
           ||||||||||| ||||||||| || ||| ||||||| ||| ||||||||||||| |
Db 126056  GACAGTCTGGCCAGGACATCGGGCGCGGGCCGCCCAGTGGTCCAAGACAAGGAGCAATTG  125997
Qy    856  cctgctgtcggctgaaggcccgttaccagacaatgagtatctgcccatggagctatat  915
           |||||||| || |||||||| ||||||||||| |||||||||||||||  |||||||
Db 125996  CCCACCGTGAGGATCAAGTGTCGTTTCCAGCACCGACCATCTGCCCATCAATGTGTAC  125937
Qy    916  aaggagttgtcagaatatgtgaccaaccactaccgacatgctgtgtccgtgctggagccc  975
           |||| |  |||||||||  |||  || |||||||| | || || |||||||||||||||
Db 125936  GGCAACTTTTTGACGTACCTCAAGGAGAACTATAAGCGCGTGTGCGAGACCCTGGAGCCG  125877
Qy    976  gccctcaatgtcaagggaacaggaggaggagtcgctagtgcactggttcacatcctgcaaagc  1035
           |||  |||||| ||||||||||||| |||||| || |||||||||||||||| |||||||
Db 125876  GTGATCGGAGTCAAGGCCAAGGACACTGGTGCTGCTGATGCACGCA  125817
Qy   1036  acaggcaaggccaaggacttcctttcagacatggccatgctggggtagaccggttcatg  1095
           ||||||||||||||||||| ||||| ||||||| || |||||||| | || |||||||
Db 125816  CAGGGATTGCGCGGCGCCTTCCTCACCCAGTGTGGTGCCCTCGATCTGCTGCGAGTTGGC  125757
Qy   1096  gagcggaacacctcatattccgcgaggaacacgtctgccactaaagccatagaagagtat  1155
           ||||| |||||| | || |||||||||||| | |||| |||| ||||||| ||||| |
Db 125756  GATCAGAGG---CTTACGTTCAGGGGCACTCCTTGGCCACCAAGAGCATGGAGGCATTC  125700
Qy   1156  atgagactgattggccagaaatacctcaaggagtgccattgggagttcatccgggctctg  1215
           ||||  ||||| |||||||||||||| |||||||| ||| ||||||||| ||||| |||
Db 125699  CTCAAGCTGACAGGCGAACAGTATCTGCAGGAGACACACTATCCCACCAATAAACGAGCTA  125640
Qy   1216  tatgaatctgaggagaactgtgaagtagacccatccaagtgca-------cagcgtcc  1266
           | ||||||||||||||||| |||| || || |||||||| ||         | |||||
Db 125639  ATTCAGTCGGAGAGGGACTGCGAGGTGGATCTCCACCAAGACCAGCGGTTCGTCGGCGGGT  125580
Qy   1267  agtctggcagagcaccaggcaacctcggatgtgctgagttggccctgtcaaggtg  1326
           || |||||| ||| ||||| ||||| ||||||| |||| || ||| || |||||
Db 125579  TCGCTGCAGCGACACGCAGGCGCCGGCCTTGGCCGCCGGCGTTCCGAGTGCATC  125520
Qy   1327  gtcaactcccattcgtgttccgagggagctgaagaggtgtttgcatcatggcggctg  1386
           ||||||| |||| ||||| ||||||||||| ||||||||||| |||| |||| |||
Db 125519  TTCGAATCGCACAAGCATTTTCCCGCCAGTTACGAAATTGCTTTGCCACGTTCCGGGAG  125460
Qy   1387  cgctgtgcagagcgggcgcggaggacaattgctgacagtctgatcagcgctgctctc  1446
           |||||||||||||||||||||| || |||||| | ||| ||||||||||| |||| |
Db 125459  CGCTTGCAGCAGCTGTGGCCCGTGGCCTGTCCAGGATATGGCCGACAACCTGATCTCGGCGACCATTTTC  125400
Qy   1447  ctggcctctctgctgccggccatcatgtcgccggccagtgtgttggactgatgcaggagtac  1506
           ||||||||||| |||||||||||||||||||||| |||||||||||||| ||||||||||
Db 125399  CTGCGCCTTCCTGCCGGCCGCCATCCTGTCGCCGGCCGCTTCAATATCACCAGCGAACTG  125340
```

Left column contig/gap annotations:

```
55834
58457:  contig of 2624 bp in length
58458
58537:  gap of unknown length
58538
60171:  contig of 1634 bp in length
60172
60251:  gap of unknown length
60252
63621:  contig of 3370 bp in length
63622
63701:  gap of unknown length
63702
66219:  contig of 2518 bp in length
66220
66299:  gap of unknown length
66300
69074:  contig of 2775 bp in length
69075
69154:  gap of unknown length
69155
72905:  contig of 3751 bp in length
72906
72985:  gap of unknown length
72986
75963:  contig of 2978 bp in length
75964
76043:  gap of unknown length
76044
78866:  contig of 2823 bp in length
78867
78946:  gap of unknown length
78947
82045:  contig of 3099 bp in length
82046
82125:  gap of unknown length
82126
85135:  contig of 3010 bp in length
85136
85215:  gap of unknown length
85216
88550:  contig of 3335 bp in length
88551
88630:  gap of unknown length
88631
92611:  contig of 3981 bp in length
92612
92691:  gap of unknown length
92692
98709:  contig of 6018 bp in length
98710
98789:  gap of unknown length
98790
104147: contig of 5358 bp in length
104148
104227: gap of unknown length
104228
111796: contig of 7569 bp in length
111797
111876: gap of unknown length
111877
118488: contig of 6612 bp in length
118489
118568: gap of unknown length
118569
126924: contig of 8356 bp in length
126925
127004: gap of unknown length
127005
137051: contig of 10047 bp in length
137052
137131: gap of unknown length
137132
148859: contig of 11728 bp in length
148860
148939: gap of unknown length
148940
149554: contig of 615 bp in length
149555
149635: gap of unknown length
150197
150198
150277: gap of unknown length
150278
150907: contig of 630 bp in length
150908
150987: gap of unknown length
150988
151646: contig of 659 bp in length
151647
151726: gap of unknown length
151727
152074: contig of 348 bp in length
152075
152154: gap of unknown length
152155
152902: contig of 748 bp in length
152903
152982: gap of unknown length
152983
153624: contig of 642 bp in length
153625
153704: gap of unknown length
153705
154242: contig of 538 bp in length
154243
154322: gap of unknown length
154323
154739: contig of 417 bp in length
154740
154819: gap of unknown length
154820
155338: contig of 519 bp in length
155339
155418: gap of unknown length
155419
155993: contig of 575 bp in length
155994
156073: gap of unknown length
156074
156806: contig of 733 bp in length
156807
156886: gap of unknown length
156887
157541: contig of 655 bp in length
157542
157621: gap of unknown length
157622
158208: contig of 587 bp in length
158209
158288: gap of unknown length
```

(contig line for 149635/150197: contig of 563 bp in length)

```
Query Match            6.0%;   Score 240.2;   DB 36;   Length 202741;
Best Local Similarity  50.6%;  Pred. NO. 8.3e-37;
Matches 748;  Conservative 0;  Mismatches 683;  Indels 48;  Gaps 5;

Qy  376  gaaaggacaaatggaatctacagagggctgtgaaacccaacaagaggacaacagc  435
         ||  | |||| ||||   || ||| | ||||||      | ||| ||||  |||
```

```
Qy   1507  ccagatgagagacctcacgaacctcaccctcatcgccaaggttatccagaacctggcc  1566
           ||| || |||||||||| ||||||| || || ||| ||| ||  || ||||||| ||||
Db 125339  CCGTCCGCACGGCCACCCGCAATCTCACACTGTGGCCAAGACCCTGCAAACATTGGCC  125280

Qy   1567  aactttccaagttacctcaaaggagagacttcctgggcttcatgaacgagtttctggag  1626
           ||  || ||| || ||| |||||| ||| ||| ||| |||| || ||| ||| || |||
Db 125279  AATTTCACCCGCTTCCAGGGCAAAGAGAACTTTATGGAGTTTTCAACGAGTTTCCTCGAG  125220

Qy   1627  ctggaatggggttctatgcagcgcaattcttgtatgagatatccaacctggacacactgac-  1685
           ||| || |||||| ||| |||||||||| || || |||||| || |||||| ||||||| ||
Db 125219  CAGGAGGCCCCTCGCATGCAACACAGTTTCTCGGAGATTATATCCACGCGGCCGGAGCACCCA  125160

Qy   1686  ---------caacagcagcagtttgagggctacatagacttgggcccgcgagctctccaca  1737
                    ||| ||| ||| |||||||||||| || |||||| | | |||| ||||
Db 125159  GCTCCAGACTCGATCCTCGATTGGGCCGGGTACATCGACCACAGGCAAACAGTTGTCCATA  125100

Qy   1738  cttcacgccctgctctggagggtgctgccccagctcagcaaggaagccctcctgaagctg  1797
           ||  || |||||   ||| ||||| || ||| ||||| |  | ||| ||| |||| |||
Db 125099  CTACACAGTTTGCTCAGCGAAAGCCTGCCAAGCTGCCGGAGGCCCAGGCAGCCACGAGCTG  125040

Qy   1798  ggccccgtcgcccccggctcctcagcgacatcagcacagcagcc  1836
           ||  ||| ||  || ||||| |||||| ||  ||| ||||||||
Db 125039  GATCCGTTGCAGCATATTCTCGATGAAATCAGCCGAGCC  125001
```

```
RESULT 15
AE003506
LOCUS       AE003506    300994 bp    DNA       INV       04-OCT-2000
DEFINITION  Drosophila melanogaster genomic scaffold 142000013386053 section 23
            of 30, complete sequence.
ACCESSION   AE003506 AE002593
VERSION     AE003506.1  GI:7293355
KEYWORDS    HTG.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 300994)
  AUTHORS   Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
            Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
            George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
            Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
            Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
            Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
            Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
            Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
            Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
            Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
            Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
            Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
            Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
            Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
            Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
            Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S.,
            Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S.,
            Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
            Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
            Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
            Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
            Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
            Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
            Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
            McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
            Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
            Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K.,
            Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
            Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
            Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
            Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
            Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R.,
            Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
            Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,
            Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
            Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
            Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
            Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
  TITLE     The genome sequence of Drosophila melanogaster
  JOURNAL   Science 287 (5461), 2185-2195 (2000)
  MEDLINE   20196006
REFERENCE   2 (bases 1 to 300994)
  AUTHORS   Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
FEATURES             Location/Qualifiers
     source          1..300994
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="X"
     mRNA            complement(join(<16198..16317,16386..16574,16635..16845,
                     16915..17053,17129..17411,17526..17607,17676..17975,
                     18107..>18483))
                     /gene="CG8527"
                     /product="CT24901"
                     /db_xref="FLYBASE:FBan0008527"
                     /db_xref="FLYBASE:FBgn0030844"
     gene            complement(<16198..>18483)
                     /gene="CG8527"
                     /map="16B4-16B6"
                     /db_xref="FLYBASE:FBan0008527"
                     /db_xref="FLYBASE:FBgn0030844"
     CDS             complement(join(16198..16317,16386..16574,16635..16845,
                     16915..17053,17129..17411,17526..17607,17676..17975,
                     18107..18483))
                     /gene="CG8527"
                     /note="CG8527 gene product"
                     /codon_start=1
                     /db_xref="FLYBASE:FBan0008527"
                     /db_xref="FLYBASE:FBgn0030844"
                     /protein_id="AAF48734.1"
                     /db_xref="GI:7293356"
                     /translation="MPQEKRHPPESRTQRFLETLVIFRRSLIYQKEFQNSTLHGVR
                     YIAESGRPIGEKFMWFCFTSIGAVTALVIMSLWEKFQTNPTITGLDTDFHQNVVFP
                     TVVCPRAAFDHDKTYEKVYNTLANVDEAQAQMYTPFLRLTSLNFENVRDAKVLSQS
                     IPQNLLDAHTIREWAFEGHIDCKNVFVSCKIRDEDIPCCDHFEPIYTEHGFCYAFNSR
                     FKSTPTEDVKTGAPHIDLYETDKKWALFFIPNSTSRIFTIFSNEEYFGSDFNAQIDWSEP
                     QLVEVRISKKNYTTDDARQLSIGQRKCIFSDEVKLNYFPDAYTFSSCMKQCRMNKAI
                     KLCKCNPPFYKPIPNVPMCSIKDFDCLDEFKSNITNIKDGLQCELSCSKTVFNIDKLI
                     KMSDRPESLGVLVEFLTWPIIRYKREVLFGWNDLLVSFGGIASLFGFSLLSGVEIIY
                     YFTLRACCMVYKNRQELYEIEEKIRQEPPPKIDLKSLKSHNPRIGDTPTSAVLKVKP
                     AEESAGPNPLGGYQRKHEKDYSLNTKNYKTPKVVPDYGYTSKTKNSWTTYDHSQYMP
                     "
     mRNA            complement(join(21273..21792,21868..21999,22058..22191,
                     22257..22823,22889..23887,24168..24320,24398..25116,
                     25191..25449))
                     /gene="CG8465"
                     /product="CT24743"
                     /db_xref="FLYBASE:FBan0008465"
                     /db_xref="FLYBASE:FBgn0030845"
     mRNA            complement(join(21273..21792,21868..21999,22058..22191,
                     22257..22823,22889..23887,24168..24320,24398..25116,
                     25191..25370,27861..28246))
                     /gene="CG8465"
                     /product="CT24757"
                     /db_xref="FLYBASE:FBan0008465"
                     /db_xref="FLYBASE:FBgn0030845"
     mRNA            complement(join(21273..21792,21868..21999,22058..22191,
                     22257..22823,22889..23887,24168..24320,24398..25116,
                     25191..25370,27861..28115,28453..28555))
                     /gene="CG8465"
                     /product="CT24733"
                     /db_xref="FLYBASE:FBan0008465"
                     /db_xref="FLYBASE:FBgn0030845"
     gene            complement(<21273..>28555)
```

CDS
```
/gene="CG8465"
/map="16B4-16B7"
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28092))
/gene="CG8465"
/note="CG8465 gene product [alt 1]"
/codon_start=1
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
/protein_id="AAF48735.1"
/db_xref="GI:7293357"
/translation="MSTYFGVYIPTSKAGCFEGSVSQCIGSIAAVNIKPSNPASGSAS
VASGSPSGSAASVQTGNADDGSAATKYEDPDYPPDSPLWLIFTEKSKALDILRHYKEA
RLREFPNLEQAESYVQFGFESIEALKRFCKAKPESKPIPIISGSGYKSSPTSTDNSCS
SSPTGNGSGFIIPLGSNSSMSNLLLSDSPTSSPSSSNVIANGRQQOMQQQOQPQQ
PDVSGEGPPFRAPTKQELVEFRKQIEGGHIDRVKRIIWENPRFLISSGDTPTSLKEGC
RYNAMHICAQVNKARIAQLLLKTISDREFTQLYVGKKGSGKMCAALNISILLDYILNMP
DKGRGETPLHFAAKNGHVAMVEVLVSYPECKSLRNHEGKEPKEIICLRNANATHVTIK
KLELLLYDPHFVPVLRSQSNTLPPKVGQFFSKPDPPNLQHKADDYEGLSVDLAISALA
GPMSREKAMNFYRRWKTPPRVSNNVMSPLAGSPFSSPVKVTPSKSIFDRSAGNSSPVH
SGRRVLFSPLAEATSSPKPTKNVPNGTNECEHNNNNVKPVYPLEFPATPIRKMKPDLF
MAYRNNSFDSPSLADDSQILDMSLSRSLNASLNDSFRERHIKNTDIEKGLEVVGRQL
ARQEQLEWREYWDFLDSFLDIGTTEGLARLEAYFLEKTEQQADKSETWNFAHLHQYF
DSMAGEQQQLRKDKNEAAGATSPSAGVMTPYTCVEKSLQVFAKRITKTLINKIGNMV
SINDTLLCELKRLKSLIVSFKDDARFISVDFSKVHSRIAHLVASYVTHSQEVSVAMRL
QLLQMLRSLRQLLADERGREQHLGCVCASLLLMLEQAPTSAVHLPDTLKTEELCCAAW
ETEQCCACLWDANLSRKTSRRKRTKSLRAAAVVQSQGLQDTSGSTGSSALHASLGVG
STSLGASRVVASASKDAWRRQQSDDEDYDSDEQVIFFDCTNVTLPYGSSSEDEENFRT
PPQSLSPGISMDLEPRYELFIFGNEPTKRDLDVLNALSNVDIKETLPHVYAWKTAME
SYSCAEMNLNVKVQKPEPWYSGTSSSHNSQPLLHPKRLLATPKLNAVVSGRRGSGPLT
APVTPRLARTPSAASIQVASEINGESVGTAVTPASPILSFAALTAATQSFQTPLNKVR
GLFSQYRDQRSYNEGDFTPLGNRN"
```

CDS
```
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28092))
/gene="CG8465"
/note="CG8465 gene product [alt 2]"
/codon_start=1
/db_xref="FLYBASE:FBan0008465"
/db_xref="FLYBASE:FBgn0030845"
/protein_id="AAF48736.1"
/db_xref="GI:7293358"
/translation="MSTYFGVYIPTSKAGCFEGSVSQCIGSIAAVNIKPSNPASGSAS
VASGSPSGSAASVQTGNADDGSAATKYEDPDYPPDSPLWLIFTEKSKALDILRHYKEA
RLREFPNLEQAESYVQFGFESIEALKRFCKAKPESKPIPIISGSGYKSSPTSTDNSCS
SSPTGNGSGFIIPLGSNSSMSNLLLSDSPTSSPSSSNVIANGRQQOMQQQOQPQQ
PDVSGEGPPFRAPTKQELVEFRKQIEGGHIDRVKRIIWENPRFLISSGDTPTSLKEGC
RYNAMHICAQVNKARIAQLLLKTISDREFTQLYVGKKGSGKMCAALNISILLDYILNMP
DKGRGETPLHFAAKNGHVAMVEVLVSYPECKSLRNHEGKEPKEIICLRNANATHVTIK
KLELLLYDPHFVPVLRSQSNTLPPKVGQFFSKPDPPNLQHKADDYEGLSVDLAISALA
GPMSREKAMNFYRRWKTPPRVSNNVMSPLAGSPFSSPVKVTPSKSIFDRSAGNSSPVH
SGRRVLFSPLAEATSSPKPTKNVPNGTNECEHNNNNVKPVYPLEFPATPIRKMKPDLF
MAYRNNSFDSPSLADDSQILDMSLSRSLNASLNDSFRERHIKNTDIEKGLEVVGRQL
ARQEQLEWREYWDFLDSFLDIGTTEGLARLEAYFLEKTEQQADKSETWNFAHLHQYF
DSMAGEQQQLRKDKNEAAGATSPSAGVMTPYTCVEKSLQVFAKRITKTLINKIGNMV
SINDTLLCELKRLKSLIVSFKDDARFISVDFSKVHSRIAHLVASYVTHSQEVSVAMRL
QLLQMLRSLRQLLADERGREQHLGCVCASLLLMLEQAPTSAVHLPDTLKTEELCCAAW
ETEQCCACLWDANLSRKTSRRKRTKSLRAAAVVQSQGLQDTSGSTGSSALHASLGVG
STSLGASRVVASASKDAWRRQQSDDEDYDSDEQVIFFDCTNVTLPYGSSSEDEENFRT
PPQSLSPGISMDLEPRYELFIFGNEPTKRDLDVLNALSNVDIKETLPHVYAWKTAME
SYSCAEMNLNVKVQKPEPWYSGTSSSHNSQPLLHPKRLLATPKLNAVVSGRRGSGPLT
APVTPRLARTPSAASIQVASEINGESVGTAVTPASPILSFAALTAATQSFQTPLNKVR
GLFSQYRDQRSYNEGDFTPLGNRN"
```

CDS
```
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25410))
/gene="CG8465"
/note="CG8465 gene product [alt 3]"
/codon_start=1
/db_xref="FLYBASE:FBan0008465"
```

```
/db_xref="FLYBASE:FBgn0030845"
/protein_id="AAF48737.1"
/db_xref="GI:7293359"
/translation="MPTHQHCHRHDGAADSPLWLIFTEKSKALDILRHYKEARLREFP
NLEQAESYVQFGFESIEALKRFCKAKPESKPIPIISGSGYKSSPTSTDNSCSSSPTGN
GSGFIIPLGSNSSMSNLLLSDSPTSSPSSSNVIANGRQQOMQQQOQPQQPDVSGE
GPPFRAPTKQELVEFRKQIEGGHIDRVKRIIWENPRFLISSGDTPTSLKEGCRYNAMH
ICAQVNKARIAQLLLKTISDREFTQLYVGKKGSGKMCAALNISILLYLNMPDKGKRGE
TPLHFAAKNGHVAMVEVLVSYPECKSLRNHEGKEPKEIICLRNANATHVTIKKLELLL
```

```
Query Match     6.0%;   Score 240.2;  DB 28;  Length 300994;
Best Local Similarity   50.6%;   Pred. No. 7.5e-37;
Matches 748;   Conservative 0;   Mismatches 683;   Indels 48;   Gaps 5;
```

```
Qy    376 gaaagggacaaatgagaatctcacagagggctggaacccaacaaggacaacagc 435
          || |||  ||| |||| | ||   ||    |||  |||          |||  |||
Db 196157 GAGCGCGCACCTTGGATCTACTCGCTGCGCAAGTCGATCGCTCCGAATGCAGAGCACACG 196216

Qy    436 cgccggtagataacgtgctgaaactatggatcatagaagctcgagagctgcccccaag 495
          |||  |||    ||| | |||  | ||| ||  | |  |   | ||   ||||  |||||
Db 196217 CGTCGCACGGACAACTCGCTGAAGATGTGGTGTACGAGGCGAAAAATCTGCCGCCCAAG 196276

Qy    496 aagcgatattactgcgagttatgcctggacgacatgctccatgcacggaccacttccaag 555
          ||||||  ||| | ||||  |||  | ||  | |||| ||  | | |  |||  || |||
Db 196277 AAGCGGTTACTTTGCCGAACTGCAATTGGACAAGAGACGCTGTACGGCCGGACTTCGGTGAAG 196336

Qy    556 cccgctcagcctcaggagacactgtctttgggcggacacttcgagttaacaactg 615
          ||     ||| |     ||||  |  |   ||||||||||| ||  |||| |||| ||
Db 196337 CTGCAGAC---------GATCTCGTGTTTGGGGGAGCACTTCGATTTCCCGACATA 196387

Qy    616 cctgctgtccgggcggctgcggctgcatctgtaccctgacctggacaaaagcggaagag 675
          |||  | ||||||  | || || |||  |   | |   |||| || |  || |  ||||||
Db 196388 CCCGAGATTAATGTGATGATCATCTTAACGTTTTCCGTGAGGTGGCACAAGAAGAAGCGG 196447

Qy    676 gacaaagcaggctacgttggcctggtgactgttccagtgccacccctggctgggccgccac 735
          ||   |||   | ||  ||  | |    ||  |||   |||   |  |   |||||  |||| |
Db 196448 GACAAATACCAATTGTGGGATCGTCGGTGAAGATACCCGTGCACGATGTCACCTCCAGATTG 196507

Qy    736 ttcacagagcatggtaccccgtgaccctgccaacaggaagtggggctctgggggtatg 795
          ||    | |||  ||  |  ||| || ||  ||   ||        ||  |||| |||
Db 196508 CCCTGCGAACAATGGTATCCCATA--------------------CTGAGCGACAAGGCGGGC 196549

Qy    796 ggctcggggggaggaggggggtcaggggcggctcaggggcaaaggaaaggaggctgt 855
          ||||  ||  || |||  |||  |||  |    |  ||||  ||||  ||  |||||||
Db 196550 GACAGTCTGGGCAGGACAGGGCGGCGGCCAGTGGGTCCAAGGACAAGGACCAATTG 196609

Qy    856 cctgctgtcggctgcaggctgaaggcccgttaccagacaatgagtatcctggcctatat 915
          |||  | |||  ||| |||  | ||  |||||  |||    | ||   | |||| |||||
Db 196610 CCCACGCTGAGGATCAAGTGTCGTTTCCAGCACCGACCATCCTGCCCATCAATGTGTAC 196669

Qy    916 aaggagtttgcagaatatgtgaccaaccactaccgcatgctgtgtcgtgtgtggagccc 975
          || |   ||  |||   ||| |    ||||     ||  ||  | || |||||  ||||
Db 196670 GGCAACTTTTGACGTACCTCAAGGAGAACTATAAGCGCGTGTCGCAGACCTGGAGCCG 196729

Qy    976 gccctcaatgtcagggcaaggagagtcgtagtgcactggttcacatcctgcaaaqc 1035
          ||||| |    |  |||||| ||| || |  || | | |||||||  | |||  ||
Db 196730 GTGATCGGAGTCAAGGCCAAGGAGGACCATTGGACAGGCACTGGTGCTGCTGTCACGCA 196789

Qy   1036 acaggcaaggccaaggactccttcagacatggccatgccctcacatggacggttcatg 1095
          ||||||||| |   |||| ||  ||||| | |||| |   | ||     |||  | |||
Db 196790 CAGGGAGTTGGCGGGCGCCTTCCTCACCGATCGGCCCTGGCCCTGAGTCGCGAGTTGGC 196849

Qy   1096 gagcggaacacctcatattccgcagaaacagctcgccactaaaagccataqaagagtat 1155
          || || |  ||||| ||    | |||||  ||| | |   || | |     || | |||
Db 196850 GATCAGAGG---CTTACGTTCAGGGGACCAACTCCTTGGCCACCAAGACATGGAGGCATTC 196906

Qy   1156 atagaactgattgccagaaatacctcaaggatgccattgggagttcatccgggctctcg 1215
          ||  ||| | |||  ||| || | | ||||  ||  || | |   |     ||| ||| |
Db 196907 CTCAAGCTGACGGCGAACAGTATCTGCAGGACACAACTATCCGCAACAACTAAACGAGCTA 196966

Qy   1216 tatgaatctgaggagaactgtaagtagacccatcaagtgca------------cagcgtcc 1266
          ||||  | ||| || | ||  ||  |  |    ||||  |||           |  ||| |
Db 196967 ATTCAGTCGGAGAGGGACTGCGGCGAGGTGGATCCCCAAGGACGACGGTTCGTCGCGCGGT 197026
```

```
Qy   1267 agtctggcagagcaccaggccaacctgcggatgtgctgtgagttggccctgtgcaaggtg 1326
               |||  ||||||||   |||  || ||||| -  ||||
Db 197027 TCGCTGCAGCGACAGCAGGCCGCCTTGCGTGGCGCGGTCCGAGGGGCTGGCAGTGCATC 197086

Qy   1327 gtcaactcccattgcgttcccgagggagctgaaggaggtgtttgcatcatgcggcctg 1386
               || - ||  ||      |||||  ||||| - | |||  ||||
Db 197087 TTCCAATCGCACAAGCATTTTCCCGCCCAGTTACGAAAATTGCTTTGCGACGTTCCGGGAG 197146

Qy   1387 cgctgtgcagagcgggccgggagagacattgctgacagcgcctgatcagcgcctcgctcttc 1446
               |||| - || - ||||  || -  ||  |  |  ||||
Db 197147 CGCTTGCAGCAGCTGGGCCGTCAGGATATGGCCGACAACCTGATCTCGCGAGCATTTTC 197206

Qy   1447 ctgcgcttcctctgccgcggccatcatgtcgccagtctgtttggactgatgcaggagtac 1506
               |||||||||||||  ||||||  ||  ||   ||   ||
Db 197207 CTGCGCTTCCTGTGCCCGGCCGCATCCTGTCGCCGCTGCTTCAATATCACCAGCGAACTG 197266

Qy   1507 ccagatgagcagacctcacgaacctcaccctcatcgccaagttatccagaacctggcc 1566
               || - ||  || - ||  ||||||  | -  |||||  |||||
Db 197267 CCGTCCCGCACGGGCCACCGCAATCTCACACTGGTGCCAAGACCCTGCAAACATTGGCC 197326

Qy   1567 aactttccaagttcacctcaaaggaggactcctgggcttcatgaacgagtttctgag 1626
               || - ||  |  |||| - |||| -  | | |||  || -  ||| |||
Db 197327 AATTTCACCCGCTTCCAGGGCAAAGAGAACTTTATGGAGTTTCTCAACGATTTCCTCGAG 197386

Qy   1627 ctggaatggggttctatgcagcaattcttgtatgagatatccaacctgacacactgac- 1685
               | |||  | - | |||||||||  | -  ||||||| -  ||  | -  ||
Db 197387 CAGGAGGGCCGCTCGCAATGCAACAGTTTCTGGAGATTATATCCACCGCGGGACCACCCA 197446

Qy   1686 ------caacagcagcagttttgagggctacaatagacttgggccgcgagctctccaca 1737
                    |||||  | | |  | |  | -  ||||| |||| -  | |||||
Db 197447 GCTCCAGAACTCGATCCTCGATTGGGCCGGGTACATCGACCAGCGCAAACAGTTGTCCATA 197506

Qy   1738 cttcacgccctgctctggaggtgtcgcccagctcagcaaggaagcccctcctgaagctg 1797
               |||| |  || - ||  ||||  |||   |  -  || -  | | ||||
Db 197507 CTACACAGTTGCTCAGCGAAAGCCTGGCCAAGCTGCCGGAGGCCAGGCCAGCGCACGAGCTG 197566

Qy   1798 ggcccgctgcccccggctcctcagccgacatcagcacagcc 1836
               | |||  | |||||||||  ||||  ||||
Db 197567 GATCCGGTTGCAGCATATTCTCGATGAAATCAGCCCGAGCC 197605
```

Search completed: January 18, 2001, 07:50:36
Job time: 22547 sec

```
              GenCore version 4.5
         Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       January 18, 2001, 11:18:13 ; Search time 231 seconds
                                            (without alignments)
                                            2777.399 Million cell updates/sec

Title:        US-09-294-298-5
Perfect score: 3981
Sequence:     1 tagggagagactgagctgcc.........tggggtgtatcccccatcct 3981

Scoring table: IDENTITY_NUC
              Gapop 10.0 , Gapext 1.0

Searched:     280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters:        561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    Issued_Patents_NA:*
              1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
              2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
              3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
              4: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
              5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.

                            SUMMARIES
```

Result No.		Score	% Query Match	Length	DB	ID	Description
1		63.6	1.6	3148	3	US-08-909-954-1	Sequence 1, Appli
2		63.4	1.6	7218	1	US-08-232-463-14	Sequence 14, Appl
3	c	62	1.6	3117	3	US-08-909-954-3	Sequence 3, Appli
4		61	1.5	2580	2	US-09-050-863-2	Sequence 2, Appli
5	c	61	1.5	5452	2	US-09-130-114-1	Sequence 1, Appli
6		61	1.5	10596	1	US-07-884-811-15	Sequence 15, Appl
7		61	1.5	10596	1	US-07-885-971-15	Sequence 15, Appl
8		61	1.5	10596	1	US-08-087-783A-15	Sequence 15, Appl
9		61	1.5	10596	1	US-08-194-088B-15	Sequence 15, Appl
10		61	1.5	10596	1	US-08-194-087-15	Sequence 15, Appl
11		61	1.5	10596	4	PCT-US93-04648-15	Sequence 15, Appl
12		61	1.5	2338	1	US-08-425-069-1	Sequence 1, Appli
13		57.6	1.4	2338	2	US-08-317-844B-1	Sequence 1, Appli
14	c	57.6	1.4	16442	3	US-08-781-891-208	Sequence 208, App
15		56	1.4	3489	1	US-08-728-323A-1	Sequence 1, Appli
16	c	54.6	1.4	32207	1	US-08-770-379-20	Sequence 20, Appl
17	c	54.4	1.4	9551	1	US-08-056-200-93	Sequence 93, Appl
18		54.4	1.4	9551	1	US-08-800-644-93	Sequence 93, Appl
19		54	1.4	543	5	5273901-6	Patent No. 5273901
20		53.6	1.3	1995	1	US-08-425-069-3	Sequence 3, Appli
21		53.6	1.3	1995	2	US-08-317-844B-3	Sequence 3, Appli
22		53	1.3	1137	5	5171843-8	Patent No. 5171843
23		52.2	1.3	1847	5	US-08-675-885-4	Sequence 4, Appli
24		51.4	1.3	697	5	5171843-10	Patent No. 5171843
25		51.4	1.3	152331	1	US-09-128-155-16	Sequence 16, Appl
26		51.2	1.3	2793	1	US-08-209-747-1	Sequence 1, Appli
27		51.2	1.3	2793	3	US-08-458-298-1	Sequence 1, Appli
28	c	51.2	1.3	51259	3	US-08-781-891-209	Sequence 209, App

29		49	1.2	2040	2	US-08-533-669A-5	Sequence 5, Appli
30		47.8	1.2	3337	1	US-08-072-610-1	Sequence 1, Appli
31		47.8	1.2	3337	3	US-08-719-822B-1	Sequence 1, Appli
32		47.6	1.2	397	3	US-09-253-691-3	Sequence 3, Appli
33		46	1.2	234	1	US-08-469-802B-3	Sequence 3, Appli
34		46	1.2	234	1	US-08-267-803B-3	Sequence 3, Appli
35		46	1.2	2818	1	US-08-366-276-1	Sequence 1, Appli
36		45	1.1	7218	1	US-08-232-463-14	Sequence 14, Appl
37		45	1.1	3456	1	US-08-190-687B-24	Sequence 24, Appl
38		45	1.1	4307	1	US-08-190-687B-7	Sequence 7, Appli
39		44.8	1.1	44377	2	US-08-804-227C-7	Sequence 7, Appli
40		44.8	1.1	44377	2	US-08-804-198-1	Sequence 1, Appli
41		44.6	1.1	533	5	5482709-5	Patent No. 5482709
42		44.4	1.1	1599	3	US-08-853-733B-1	Sequence 1, Appli
43		44.4	1.1	1603	3	US-08-675-885-6	Sequence 6, Appli
44		44.4	1.1	1931	2	US-09-130-114-2	Sequence 2, Appli
45	c	43.8	1.1	2385	2	US-08-393-333-1	Sequence 1, Appli

```
                            ALIGNMENTS

RESULT 1
US-08-909-954-1
; Sequence 1, Application US/08909954A
; Patent No. 6100058
; GENERAL INFORMATION:
;  APPLICANT: Allen, Maxine J.
;  APPLICANT: Buckler, Alan J.
;  TITLE OF INVENTION: GAP12 Genes and their Uses
;  FILE REFERENCE: SEQ-11P
;  CURRENT APPLICATION NUMBER: US/08/909,954A
;  CURRENT FILING DATE: 1997-08-12
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;  LENGTH: 3148
;  TYPE: DNA
;  ORGANISM: H. sapiens
US-08-909-954-1

Query Match             1.6%;  Score 63.6;  DB 3;  Length 3148;
Best Local Similarity   50.0%;  Pred. No. 4.8e-06;
Matches  159;  Conservative   0;  Mismatches  159;  Indels   0;  Gaps   0;

Qy   1261 gcgtccagtctggcagagcaccaggccaactggcggatgtcggattggctgagttggccctgtgc 1320
          |  ||||| ||||||| | ||||||| || | ||||||||||  || ||| ||||| |  || ||
Db   1518 gagaccagcctggggctgctgacgggctacctggggccatcgtggacggaccgccatcgtgggc 1577

Qy   1321 aaggtgtcaactcccattgcctgttcccgaggagagctgaaggaggtgtttgtcatcatgg 1380
          ||||  ||| || |||| ||  ||  | || | |||| |||| ||| |  |     |   |
Db   1578 tccgtggggcgctgaccccgccatgccatggccagcctccaagcagctgacagctcaccggggagtg 1637

Qy   1381 cggctgcgctgtgcagagcgggcgcgggggagggacattgctgacaggctgatcagcgcctcg 1440
          || ||| |||||||||||  ||| ||| |||||| | | | ||||||| ||| |||| |||||
Db   1638 gaggagcgcttcccagctcctgcccgcccagccccagggcaccaggatgtgaagtaccctggccatcagtggattt 1697

Qy   1441 ctcttcctgcgcttcctcctgcccggccatcatgtcgccagtctgtttggactgatgcag 1500
          ||||||  || ||   ||| ||  ||  || | |  |||||||| || | | ||||||||
Db   1698 ctcttcttgcgattcttcgcacctgcacctccaaagctgtttgacctttcgggac 1757

Qy   1501 gagtaccagatgagcagacagacctcacgaacctcacccctcatcgccaaggttatccagaac 1560
          |||| ||  ||||||  |||||| |||||| |  | |||||| | |||| || ||||||||
Db   1758 caacacgcgaccccccagactagcccagactagcgctcactgctgttgctgctgccaaggctgtgcagagc 1817

Qy   1561 ctggccaacttttccaag 1578
          |||||||| || ||||||
Db   1818 attggaaacctgggccag 1835

RESULT 2
US-08-232-463-14/c
```

```
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
;   APPLICANT: DORNER, F.
;   APPLICANT: SCHEIFLINGER, F.
;   APPLICANT: FALKNER, F. G.
;   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;   NUMBER OF SEQUENCES: 52
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Foley & Lardner
;     STREET: 1800 Diagonal Road, Suite 500
;     CITY: Alexandria
;     STATE: VA
;     COUNTRY: USA
;     ZIP: 22313-0299
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/232,463
;     FILING DATE:
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/07/935,313
;     FILING DATE:
;     APPLICATION NUMBER: EP 91 114 300.6
;     FILING DATE: 26-AUG-1991
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BENT, Stephen A.
;     REGISTRATION NUMBER: 29,768
;     REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (703)836-9300
;     TELEFAX: (703)683-4109
;     TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 7218 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;     CLONE: pTzgpt-F1s
US-08-232-463-14

Query Match               1.6%;  Score 63.4; DB 1;  Length 7218;
Best Local Similarity     3.3%;  Pred. No. 7.4e-06;
Matches    13;  Conservative  233;  Mismatches 149;  Indels    0;  Gaps    0:

Qy  3164 acatcgagagtgcacacattgagcggaagagtacaaagctgaaggagtactcgaagtcca 3223
              : :::: ::  ::  ::::::::  :::  :: :::::: ::  ::::
Db  1436 ACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1377

Qy  3224 tggacagagaccgactggacaggtgaaggagtacgaggagagatccactcactgaagg 3283
         :::::::::::::::::::::::::::::::::::::::::::::::::::
Db  1376 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1317

Qy  3284 aaaggctacacatgtccaaccggaagctgaagagtacgagcggagcggaggctgctgtcccagg 3343
         :: ::: : :::  :::  :::  :::::::  :: :  ::
Db  1316 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1257

Qy  3344 aagagcagaccagcagaagatcctgatgcagtaccaagcccgcctggagcagagcgagaagc 3403
         : ::::  :: ::  :::  :::  :::::: :: :::  ::
Db  1256 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1197

Qy  3404 gcttgaggcagcagcaggtgagaagaggactcccagatcaagagcatcattggcaggctga 3463
         : ::: ::::::  ::  :::: :::  ::::  :::: :: ::::
Db  1196 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1137

Qy  3464 tgctgtggaggaggagctgcgccggaccaccccgaccatggcctgagccgctgcctgaac 3523
```

```
Db  1136 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1077
              :  : : ::::::::: : :::     ::  :   ::: :::   : :: :::
Qy  3524 ccaagaagaggctgcgacgctcagagaggagcagc 3558
              :: :: ::   :::  : :::::    ::::::
Db  1076 RRRRRRRRRRRATCGCAAGTCCCTGCAAGCTCCCTGCACCTGCAGC 1042


RESULT   3
US-08-909-954-3
; Sequence 3, Application US/08909954A
; Patent No. 6100058
; GENERAL INFORMATION:
;   APPLICANT: Allen, Maxine J.
;   APPLICANT: Buckler, Alan J.
;   TITLE OF INVENTION: GAP12 Genes and their Uses
;   FILE REFERENCE: SEQ-11P
;   CURRENT APPLICATION NUMBER: US/08/909,954A
;   CURRENT FILING DATE: 1997-08-12
;   NUMBER OF SEQ ID NOS: 15
;   SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 3117
;   TYPE: DNA
;   ORGANISM: M. musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (0)...(0)
US-08-909-954-3

Query Match               1.6%;  Score 62; DB 3;  Length 3117;
Best Local Similarity    49.7%;  Pred. No. 1.1e-05;
Matches   158;  Conservative   0;  Mismatches 160;  Indels    0;  Gaps    0:

Qy  1261 gcgtccagtctggcagagcagaccacaggccaactgccggatgtgtgctgtgagttggccctgtgc 1320
         ||| ||||| |||| ||||     || || ||  | | || |  ||| |  || |
Db  1347 gagaccagcctggggctgctgaccggataccaggtccgttggacgccatcgtgagc 1406

Qy  1321 aagtggtcaactccattgcgtgttccgagggagctgaagaggggtttgcatcatgg 1380
         |  ||  |  |    ||| || |||  |   ||| |||
Db  1407 tctacaggggctgctgccactgccttggccttaagcagcagctccagcggtgtg 1466

Qy  1381 cggctgcgtgtgcagagcgggggcggagggacattgctgacaggctgatcagcgcctcg 1440
         | ||   |||| || |   |  | ||||   | |  ||| ||   | |
Db  1467 gagaagcgcttctctgggatagagcatcaggatgtgaagtacctggcatcagtggcttc 1526

Qy  1441 ctcttcctgcgttcctctcgcccggccatcatgtcgccggccagtcgtcttggactgatgcag 1500
         ||||||||||| ||| | |  ||||||     ||||| ||  | |  ||||||  |  |
Db  1527 ctctttctgcggtcttttgcgcctgagcatcctcacaccgaaactgtttgacctcagagac 1586

Qy  1501 gagtaccagatgagcgagacctcacgaacctcaccctcatcgccaagttatccagaac 1560
         || ||||||  | ||||  |||||||  || |   |||||| ||| | |||||
Db  1587 cagcacgcagaccccagacccagacccgttccctgctccgctcgccaaggctgtgcagagc 1646

Qy  1561 ctggccaactttccaag 1578
         | |  | |  ||  ||
Db  1647 attgggaacctgggccag 1664


RESULT   4
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
;   APPLICANT: Lao, Ying
;   APPLICANT: Hiang, Betty
;   APPLICANT: Payan, Don
;   TITLE OF INVENTION: Mammalian Protein Interaction Cloning
;   TITLE OF INVENTION: System
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
```

Copyright GenCore version 4.5
(c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 20:02:07 ; Search time 505.11 Seconds
                                       (without alignments)
                                       2960.767 Million cell updates/sec

Title: US-09-294-298-5
Perfect score: 3981
Sequence: 1 tagggagagactgagctgcc.........tggggtgctatcccatcct 3981

Scoring table: IDENTITY_NUC
                  Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_36:*
    1: /SIDS6/gcgdata/geneseq/geneseqn/NA1980.DAT:*
    2: /SIDS6/gcgdata/geneseq/geneseqn/NA1981.DAT:*
    3: /SIDS6/gcgdata/geneseq/geneseqn/NA1982.DAT:*
    4: /SIDS6/gcgdata/geneseq/geneseqn/NA1983.DAT:*
    5: /SIDS6/gcgdata/geneseq/geneseqn/NA1984.DAT:*
    6: /SIDS6/gcgdata/geneseq/geneseqn/NA1985.DAT:*
    7: /SIDS6/gcgdata/geneseq/geneseqn/NA1986.DAT:*
    8: /SIDS6/gcgdata/geneseq/geneseqn/NA1987.DAT:*
    9: /SIDS6/gcgdata/geneseq/geneseqn/NA1988.DAT:*
   10: /SIDS6/gcgdata/geneseq/geneseqn/NA1989.DAT:*
   11: /SIDS6/gcgdata/geneseq/geneseqn/NA1990.DAT:*
   12: /SIDS6/gcgdata/geneseq/geneseqn/NA1991.DAT:*
   13: /SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:*
   14: /SIDS6/gcgdata/geneseq/geneseqn/NA1993.DAT:*
   15: /SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:*
   16: /SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:*
   17: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:*
   18: /SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:*
   19: /SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:*
   20: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:*
   21: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		95.8	2.4	2647	18	T76739	cDNA encoding acti
2		63.6	1.6	1337	20	Z17263	Human gene express
3		61	1.5	795	19	V55830	FLGA insert stabil
4		61	1.5	799	19	V55831	Nucleotide sequenc
5	c	61	1.5	5452	20	X90923	Anti-sense strand
6	c	61	1.5	8705	20	Z23778	Vector pShuttle DN
7		61	1.5	9600	19	V21683	Vector plasmid pCM
8		61	1.5	10380	20	Z22248	Nucleotide sequenc
9		61	1.5	10596	14	Q51731	Plasmid pCisEBON f
10		61	1.5	10596	17	T40348	Plasmid pCisEBON f
11		61	1.5	10596	20	X15650	Nucleotide sequenc
12		59.6	1.5	1925	20	X90924	Epstein Barr Virus

13		59.4	1.5	1908	8	N71064	Gene encoding Plas
14	c	59.2	1.5	1000	21	A02484	Human colon cancer
15	c	57.8	1.5	49999	20	Z23895	Murine LOBO homolo
16		57.6	1.4	2338	19	V23249	Nephila clavipes s
17		57.6	1.4	2338	21	Z38195	N. clavipes spider
18		56	1.4	2000	8	N71065	Gene encoding Plas
19		56	1.4	2338	12	Q14183	N.clavipes draglin
20	c	56	1.4	16442	18	X83006	Partial mouse WRN
21		55.8	1.4	2004	18	T85356	Nephila clavipes s
22	c	54.6	1.4	32207	20	V73805	KSHV LUR DNA (nucl
23	c	54.6	1.4	137507	19	V19941	KSHV long unique c
24		54.4	1.4	9551	20	Z22301	cDNA encoding a hu
25		54	1.4	543	13	Q23092	Antigen tc-7a gene
26		53.6	1.3	1995	12	Q14184	N.clavipes draglin
27		53.6	1.3	1995	19	V23250	Nephila clavipes s
28		53.6	1.3	1995	21	Z38196	N. clavipes spider
29		53.4	1.3	4055	20	Z40484	Human ZC2 DNA. Ho
30		53.4	1.3	4180	21	A10669	Human protein kina
31		53	1.3	1137	20	Q33061	Plasmodium vivax c
32	c	52.4	1.3	114955	20	X53491	Human adenosine A1
33		52.2	1.3	1847	19	V04694	Cell cycle protein
34		51.4	1.3	3198	20	X02974	Human IL-1ra BAC c
35		51.2	1.3	2744	16	Q98470	MiSP1-containing p
36	c	51.2	1.3	51259	18	X83007	Partial mouse WRN
37	c	51	1.3	10732	21	A10594	Gene encoding a su
38	c	50.2	1.3	1686	16	Q87587	DNA encoding Leuco
39		49	1.2	2040	18	T62137	Leishmania brazili
40		49	1.2	2040	19	V47557	Leishmania antigen
41	c	48.8	1.2	1218	21	A02488	Human colon cancer
42		48.8	1.2	2313	20	Z07197	Human lung tumour
43		48.6	1.2	1432	19	V47585	Leishmania antigen
44		47.8	1.2	3337	17	T34620	P. vivax ESP-1 blo
45		47.8	1.2	3337	20	X15174	DNA encoding a sec

ALIGNMENTS

RESULT   1
T76739
ID   T76739 standard; cDNA; 2647 BP.
XX
AC   T76739;
XX
DT   08-OCT-1997 (first entry)
XX
DE   cDNA encoding active type R-Ras protein binding protein p98.
XX
KW   Active-type-R-Ras protein; p98; tumour; disease; cancer;
KW   cellular Ca ion; phospholipid; tumour forming promoter;
KW   tumour forming inhibitor; ss.
XX
OS   Bos taurus.
XX
FH   Key   Location/Qualifiers
FT   CDS   13..2517
FT   /*tag= a
FT   /product= p98
XX
PN   JP09135688-A.
XX
PD   27-MAY-1997.
XX
PF   16-NOV-1995;   95JP-0298720.
XX
PR   16-NOV-1995;   95JP-0298720.
XX
PA   (KIRI) KIRIN BREWERY KK.
XX
DR   WPI; 1997-335993/31.
DR   P-PSDB; W24227.
XX
PT   Active R-Ras protein-binding protein p98 - used in the treatment of

```
PT    cancer, and diseases related to cellular calcium ion concentrations
XX
PS    Claim 9; Page 9-12; 18pp; Japanese.
XX
CC    This sequence encodes a protein capable of binding to active-type-R-Ras
CC    protein. The protein has a molecular weight of 98 kD by SDS-PAGE, and
CC    is designated p98. The protein coding sequence can be inserted into a
CC    suitable vector to treat certain tumours or diseases (various cancers)
CC    in which active R-Ras protein involved.  Because active R-Ras protein
CC    is involved in formation and inhibition of tumours and related to
CC    diseases in which cellular Ca ions or phospholipids are involved, the
CC    screening method can be used for screening e.g. a tumour forming promoter
CC    or tumour forming inhibitor.
XX
SQ    Sequence 2647 BP; 633 A; 764 C; 769 G; 481 T; 0 other;

Query Match             2.4%;  Score 95.8;  DB 18;  Length 2647;
Best Local Similarity  50.9%;  Pred. No. 1.5e-12;
Matches  310;  Conservative   0;  Mismatches  287;  Indels   12;  Gaps    3;

Qy    987 caagggcaaggagaggagtgctagtgcactggttcacatcctgcaaagcacaggcaaggc 1046
          ||||||| ||||  ||| ||| ||||||  | ||||||||||||||| || |||||||||
Db    996 cagggaagcaggaggaggcggcatccgtcggtgcggctcttcctgccattacggccgggt 1055

Qy   1047 caaggacttccttcagacatggccatgtcagaggtagaccgttcatggacgggaaca 1106
          ||||| ||||||||||||||  || ||||||| || ||||||  | || |
Db   1056 ggtgccctcatcagcgaccgccatcgccagggtccgcaggaccgaggtccgcaggacc---ccaa 1112

Qy   1107 cctcatattccgcgagaacacgctccgccactaaagccatagaagagtatatgagactgat 1166
          || |||||| ||||||| ||| |||| |  | |||||| ||  ||| |||
Db   1113 tacgatcttccgagggaactgctgcgacatccagtgcatcgatcagacgatgaagctggc 1172

Qy   1167 tggccagaaatacccaaggatgccattggggagtcatccggctcgtgtatgaatctga 1226
          ||| | ||||||||| ||| |||||| |||||||  ||  ||| || | || ||||
Db   1173 aggcatgcagtacctacacgtcacctgaaaccccacatagagaagatttgccagagtca 1232

Qy   1227 ggagaactgaagtagacccatcaagtgcacagcgtcc---agtctggcagagcacca 1283
          || |||||||| || ||||||| |||||||   ||     || ||| ||||||||| ||
Db   1233 taagtcctgtgaaattgaccccgtgaggctgaaggacgacggcgagagctagagagcaacat 1292

Qy   1284 ggccaaactgcgagtgtgctgatgttgccctgtgcaaggtggtcaactccattgcgt 1343
          |||||||||||| ||||||||||||| || ||||||||| |||||| || || |||
Db   1293 ggagaaactgcgcagtttgggaccgcagtttggaccggttcagcgtcatccaccaaatcggggtgag 1352

Qy   1344 gttcccgggagctgaaggaggtgttgcatcatgcggctgcgctgtgcagagcgggg 1403
          ||||| || ||| || || ||| |||||| |  |||  ||||||| ||||||  ||||
Db   1353 ctgcccacgccatgtgtgacatcttctcctgcgcggaggcagcgccaagcgctt 1412

Qy   1404 ccgggagcacattgctg-----acaggctgatcagcgccctcgcctcttcctcgcgttcct 1457
          || |||||| || ||| |           || |||||||||||||||||||| ||||
Db   1413 ccaggatgacccggacgtcgcggtacacagctgagcagcttcattttctgcggttctt 1472

Qy   1458 ctgccggccatcatgtcgccccagtctgtttggactgatgcaggagtaccccagatggagca 1517
          || ||||||||||| |||| ||||||| |||||||||||| ||| ||||||||| ||||
Db   1473 cgctccagccatcctgtccccaaccttccagctcacgccccaccacacggaccccaca 1532

Qy   1518 gacctcacgaaccctcaccctcatcgcacaaggttatccagaacctggccaacttttccaa 1577
          || ||| || ||||| ||||||  |||||||| || ||  ||||| |||||  |||||||
Db   1533 gacgtcgaggacggcgacacttgtttcgaagaccattcagacctctgggcagcctgtccaa 1592

Qy   1578 gttacctc 1586
          || || ||
Db   1593 gtccaagtc 1601

RESULT   2
Z17263
ID    Z17263 standard; cDNA; 1337 BP.
XX
AC  • Z17263;
XX
DT    12-OCT-1999 (first entry)
```

```
XX
DE    Human gene expression product cDNA sequence SEQ ID NO:4735.
XX
KW    Human; gene; gene expression product; diagnosis; therapy; probe;
KW    detection; mapping; tissue typing; profiling; forensic; cancer;
KW    genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
XX
OS    Homo sapiens.
XX
PN    WO993972-A2.
XX
PD    05-AUG-1999.
XX
PF    28-JAN-1999;       99WO-US01619.
XX
PR    03-APR-1998;       98US-0080666.
PR    28-JAN-1998;       98US-0072910.
PR    24-FEB-1998;       98US-0075954.
PR    31-MAR-1998;       98US-0080114.
PR    03-APR-1998;       98US-0080515.
XX
PA    (CHIR ) CHIRON CORP.
PA    (HYSE-) HYSEQ INC.
XX
PI    Crkvenjakov R,  Dickson M,  Drmanac R,  Drmanac S;
PI    Escobedo J,  Garcia PD,  Garcia V,  Giese K,  Innis MA;
PI    Jones WL,  Kassam A,  Kennedy GC,  Kita D,  Labat I;
PI    Lamson G,  Leshkowitz D,  Pot D,  Randazzo F,  Reinhard C;
PI    Stache-Crain B,  Sudduth-Klinger J,  Williams LT;
XX
DR    WPI; 1999-494092/41.
XX
PT    Novel human genes and their expression products which are
PT    differentially expressed in different cell types
XX
PS    Claim 1; Page 2250-2251; 2479pp; English.
XX
CC    The present invention describes a library of human polynucleotides
CC    comprising the sequences given in z12532 to z17779. Also described is a
CC    method of detecting differentially expressed genes correlated with the
CC    cancerous state of a mammalian cell, comprising detecting at least one
CC    differentially expressed gene product in a test sample from a cell
CC    suspected of being cancerous, where the gene product is encoded by one
CC    of the 5248 polynucleotide sequences given in z12532 to z17779. The
CC    polynucleotides can be used as a source of primers and probes, which can
CC    be used for a variety of purpose, e.g. detection of expression levels,
CC    mapping, tissue typing or profiling, forensics, genetic analysis and
CC    detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC    can be used for raising antibodies for experimental, diagnostic and
CC    therapeutic purposes. The polynucleotides may also be used to construct
CC    arrays for diagnostics (which may be used to determine function of an
CC    encoded protein); and to detect differences in expression levels between
CC    two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC    identify a genetic predisposition or susceptibility to a disease such as
CC    cancer). The polynucleotides of the invention are especially used in the
CC    diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC    and lung cancer. The polynucleotides can also be used to screen for
CC    peptide analogues and antagonists.
XX
SQ    Sequence 1337 BP; 42 A; 577 C; 27 G; 22 T; 669 other;

Query Match             1.6%;  Score 63.6;  DB 20;  Length 1337;
Best Local Similarity  33.0%;  Pred. No. 2.5e-05;
Matches  147;  Conservative   0;  Mismatches  298;  Indels   0;  Gaps    0;

Qy   3535 ctgctcgacgtcagagaggagcagcttcccccctgggtccaacaaccgcgtgtgacgc 3594
          | |||||  |||||  ||||||||||||||||||||||||||||nnnnnnnnc
Db    419 cnncnccccnnnnccnnnnncccnnnnnnnnncccccccccccnnnnnnnnnnnnnc 478

Qy   3595 tggcccacctggaacggcctggccccccaccccggccccaccccggctgcaga 3654
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 nnccccnncnncnccccccnncccccccccccccccccccnnnnccccnnccncc 538
```

`:` `..`

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 07:58:49 ; Search time 6516.34 Seconds
                                             (without alignments)
                                             4281.055 Million cell updates/sec

Title: US-09-294-298-5
Perfect score: 3981
Sequence: 1 taggagagactgagctgcc..........tggggtgctatcccccatcct 3981

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 791742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*

```
  1: gb_est1:*           44: em_esthum2:*
  2: gb_est2:*           45: em_esthum3:*
  3: gb_est3:*           46: em_esthum4:*
  4: gb_est4:*           47: em_esthum5:*
  5: gb_est5:*           48: em_esthum6:*
  6: gb_est6:*           49: em_esthum7:*
  7: gb_est7:*           50: em_esthum8:*
  8: gb_est8:*           51: em_esthum9:*
  9: gb_est9:*           52: em_esthum10:*
 10: gb_est10:*          53: em_esthum11:*
 11: gb_est11:*          54: em_esthum12:*
 12: gb_est12:*          55: em_esthum13:*
 13: gb_est13:*          56: em_esthum14:*
 14: gb_est14:*          57: em_esthum15:*
 15: gb_est15:*          58: em_esthum16:*
 16: gb_est16:*          59: em_esthum17:*
 17: gb_est17:*          60: em_esthum18:*
 18: gb_est18:*          61: em_esthum19:*
 19: gb_est19:*          62: em_esthum20:*
 20: gb_est20:*          63: em_estin1:*
 21: gb_est21:*          64: em_estin2:*
 22: gb_est22:*          65: em_estin3:*
 23: gb_est23:*          66: em_estin4:*
 24: gb_est24:*          67: em_estov1:*
 25: gb_est25:*          68: em_estov2:*
 26: gb_est26:*          69: em_estpl1:*
 27: gb_est27:*          70: em_estpl2:*
 28: gb_est28:*          71: em_estpl3:*
 29: gb_est29:*          72: em_estpl4:*
 30: gb_est30:*          73: em_estpl5:*
 31: gb_est31:*          74: em_estro1:*
 32: gb_est32:*          75: em_estro2:*
 33: gb_est33:*          76: em_estro3:*
 34: gb_est34:*          77: em_estro4:*
 35: gb_est35:*          78: em_estro5:*
 36: gb_est36:*          79: em_estro6:*
 37: gb_est37:*          80: em_estro7:*
 38: gb_est38:*          81: em_estro8:*
 39: gb_est39:*          82: em_estro9:*
 40: gb_est40:*          83: em_estro10:*
 41: em_estba:*          84: em_estro11:*
 42: em_estfun:*         85: em_estro12:*
 43: em_esthum1:*        86: em_estro13:*
                         87: gb_est41:*
                         88: gb_est42:*
                         89: gb_est43:*
                         90: gb_est44:*
                         91: gb_est45:*
                         92: gb_est46:*
                         93: gb_est47:*
                         94: gb_est48:*
                         95: gb_est49:*
                         96: gb_est50:*
                         97: gb_est51:*
                         98: gb_est52:*
                         99: gb_est53:*
                        100: gb_est54:*
                        101: gb_est55:*
                        102: gb_est56:*
                        103: gb_est57:*
                        104: gb_est67:*
                        105: gb_est68:*
                        106: gb_est69:*
                        107: gb_est70:*
                        108: gb_est71:*
                        109: gb_est72:*
                        110: gb_est73:*
                        111: gb_est74:*
                        112: em_esthum21:*
                        113: em_esthum22:*
                        114: em_esthum23:*
                        115: em_estom1:*
                        116: em_estom2:*
```

```
117: em_estpl6:*
118: em_estpl7:*
119: em_estpl8:*
120: em_estrol4:*
121: em_estrol5:*
122: em_estrol6:*
123: em_estrol7:*
124: em_estrol8:*
125: em_estrol9:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est75:*
136: gb_est76:*
137: gb_est77:*
138: gb_est78:*
139: gb_est79:*
140: gb_est80:*
141: gb_est81:*
142: gb_est82:*
143: gb_est83:*
144: gb_est84:*
145: gb_est85:*
146: gb_est86:*
147: gb_est87:*
148: gb_gss1:*
149: gb_gss2:*
150: gb_gss3:*
151: gb_gss4:*
152: em_gss1:*
153: em_gss2:*
154: em_gss3:*
155: gb_gss4:*
156: gb_gss5:*
157: gb_gss6:*
158: gb_gss7:*
159: gb_gss8:*
160: gb_gss9:*
161: em_gss5:*
162: em_gss6:*
163: em_gss7:*
164: em_gss8:*
165: em_gss9:*
166: em_gss10:*
167: em_gss11:*
168: gb_gss10:*
169: gb_gss11:*
170: em_gss12:*
171: gb_gss12:*
172: gb_gss13:*
173: gb_gss14:*
174: gb_gss15:*
175: gb_gss16:*
176: gb_gss17:*
177: em_gss17:*
178: gb_gss18:*
179: em_gss13:*
180: gb_gss19:*
181: gb_gss20:*
182: gb_gss21:*
183: gb_gss22:*
184: gb_gss23:*
185: em_gss14:*
186: em_gss15:*
187: em_gss16:*
188: em_gss17:*
189: em_gss18:*

190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	490.4	12.3	571	37	AV590698	AV590698 AV590698
	2	383.8	9.6	474	181	AZ399131	AZ399131 1M0164P14
	3	294.6	7.4	593	95	AW779747	AW779747 hn85f10.x
	4	250.4	6.3	463	108	BE463433	BE463433 hw23d06.x
	5	244.6	6.1	427	11	AA709010	AA709010 zf94h11.s
	6	233.6	5.9	597	90	AW421227	AW421227 fj94f10.x
	7	228.8	5.7	682	173	AQ993964	AQ993964 RPCI-23-2
c	8	222	5.6	566	22	AI606711	AI606711 m158d01.y
c	9	220.2	5.5	413	10	AA691827	AA691827 vt05d01.r
	10	214	5.4	707	28	AJ396547	AJ396547 AJ396547
	11	210.2	5.3	555	27	AI958829	AI958829 fd22g01.y
c	12	201	5.0	628	108	BE490923	BE490923 db38g08.x
	13	197.2	5.0	516	111	BE695645	BE695645 MR1-BT080
	14	195.8	4.9	347	87	AW204612	AW204612 UI-H-BI1-
	15	195.4	4.9	596	136	BE848541	BE848541 uw39f05.y
c	16	190	4.8	481	135	BE773015	BE773015 RC1-FT013
	17	185	4.6	542	93	AW656799	AW656799 109245 MA
	18	180.6	4.5	296	23	AI650331	AI650331 wa18f01.x
	19	177.6	4.5	345	39	AW076911	AW076911 fj03d08.y
c	20	173	4.3	511	22	AI577567	AI577567 UI-R-Y0-v
	21	168.8	4.2	388	183	B68206	B68206 CIT978SK-A-
c	22	165.6	4.2	474	87	AW205989	AW205989 UI-H-BI1-
c	23	161.2	4.0	460	22	AI609604	AI609604 tw91e09.x
c	24	157	3.9	449	9	AA554056	AA554056 n101g07.s
	25	156.2	3.9	1020	191	CNS02VGW	AL215753 Tetraodon
	26	154.6	3.9	440	2	AA107246	AA107246 m158d01.r
c	27	149.4	3.8	651	19	AI327335	AI327335 mp74c11.x
	28	148.6	3.7	877	192	CNS04DL0	AL285885 Tetraodon
c	29	146.8	3.7	443	40	AW136165	AW136165 UI-H-BI1-
c	30	146.8	3.7	464	95	AW826497	AW826497 fk62h05.x
	31	146	3.7	452	9	AA554055	AA554055 n101g06.s
	32	145.4	3.7	407	133	BE015978	BE015978 fk62h05.y
c	33	139.8	3.5	382	22	AI562171	AI562171 vv73f04.x
c	34	138.6	3.5	317	91	AW527364	AW527364 UI-R-BOI-
	35	138.6	3.5	457	2	AA118397	AA118397 mp74c11.r
	36	138.4	3.5	268	92	AW593427	AW593427 hg15e10.x
	37	136.6	3.4	379	9	AA575948	AA575948 nm56e05.s
	38	135.6	3.4	554	111	BE723198	BE723198 192621 MA
	39	133.4	3.4	336	26	AI874961	AI874961 u127f03.x
	40	131.8	3.3	694	28	AJ399084	AJ399084 AJ399084
c	41	130.4	3.3	335	134	BE085114	BE085114 CM2-BT066
c	42	128.4	3.2	1006	190	CNS00HWF	AL073826 Drosophil
	43	128.2	3.2	631	191	CNS01XQ5	AL172022 Tetraodon
	44	125.2	3.1	975	191	CNS03ANH	AL235430 Tetraodon
c	45	124.2	3.1	397	111	BE695651	BE695651 MR1-BT080

ALIGNMENTS

```
RESULT  1
AV590698
LOCUS       AV590698         571 bp    mRNA                 EST       29-AUG-2000
DEFINITION  AV590698 Bos taurus brain fetus Bos taurus cDNA clone E1BR014B04
            5', mRNA sequence.
ACCESSION   AV590698
VERSION     AV590698.1  GI:9701691
KEYWORDS    EST.
SOURCE      cow.
```

ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 571)
AUTHORS   Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
            Suzuki,H.
TITLE   bovine cDNA sequencing
JOURNAL   Unpublished (2000)
COMMENT   Contact: Yoshikazu Suginoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES   Location/Qualifiers
     source   1..571
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="E1BR014B04"
                     /clone_lib="Bos taurus brain fetus"
                     /tissue_type="brain"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /note="Vector: pZL1; Site_1: Sall; Site_2: NotI; Poly A
                     was deleted from a NotI site"
BASE COUNT   124 a   161 c   169 g   116 t   1 others
ORIGIN

Query Match   12.3%;   Score 490.4;   DB 37;   Length 571;
Best Local Similarity 93.2%;   Pred. No. 3.8e-104;
Matches 523;   Conservative   0; Mismatches 37;   Indels   1; Gaps   1;

Qy   992 gcaaggaggaggtcgctagtgccactggttcacatcctgcaaagcacaggccaaggccaagg 1051
          ||||||||||||| || ||||||||||||||| |||| || | |||||||| |||||||||
Db   11 GCAAGGAGGAGGTCGCCCAGTGCTCTGGTTCACATCCTGCAGAGTACAGGCCAAGGCCAAGG 70

Qy   1052 acttcctttcagacatgccatgtcagaggtagaccggttcatggagcggaacacctca 1111
          |||||||||||||||||| ||||||| |||| ||||||| || |||||||||||||||
Db   71 ACTTCCTTTCAGACATGGCCATGTCCGAGGTGACCGGTTCATGGAACGAGAACACCTCA 130

Qy   1112 tattccggagaacacgctcgccactaaagccatagaaagagtatatgagacttgattggcc 1171
          |||||||||||||||| ||||||||||||||||| |||||||||| |||||||| ||||||
Db   131 TATTCCGGAGAACACGCTCGCCACTAAAGCCATAGAAGAGTACATGAGACTGATTGGTC 190

Qy   1172 agaaatacctcaaggatgccattgggagttcatccggctctgtatgaatctgaggaga 1231
          |||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db   191 AGAAATACCTCAAGGATGCCATCGGGGAATTCATCCGTGCTCGTGTATGAATCTGAGGAGA 250

Qy   1232 actgtgaagtagacccatcaagtgcacagcgtccagtctggcagagcaccaggccaacc 1291
          ||||||| || |||||||||||||||| |||| |||| |||||| |||||||||||||||
Db   251 ACTGTGAGGTGGACCCCATCAAGTGCACGGCCCCATCAAGTTGGCTGAGCACCAGGCCAACC 310

Qy   1292 tgcggatgtgctgagttggccctgcaaggtggtcaactccattgcgtgttcccga 1351
          |||||| ||||||||||||||||||||||||||| ||||||||| |||||||||||
Db   311 TGCGGATGTGCTGAGTTGGCCCTGCAAGGTGGTCAACTCCCATTGCGTGTTCCCGA 370

Qy   1352 gggagctgaaggaggtgtttgcatcatggccgctgcgctgtgcagagcgggcggagg 1411
          |||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db   371 GGGAGCTGAAGGAGGTGTTTGCCGTCCTGGCGACTGCGCTGTGCAGAGCGGGGCCGAGAGG 430

Qy   1412 acattgctgacagctcatcacgcctcgccttcctcgcgctcctctcgcccggccatca 1471
          ||||||| ||||||||||||||||||| |||||||||||| |||||||||||||||||
Db   431 ACATTGCCGACACCGGCGTGATCACGGCCCTCGCTTCCTCGCGCTCCTCTCCGCCCGGCCATTA 490

Qy   1472 tgtcgcccagtctgtt·tggactgatgcaggagtacccagatgagacctcacgaacc 1530
          |||||||||||| || ||||||||||||||||||||||||||||||||||||| |||
Db   491 TGTCGCCCAGCCTNTTCGGGGCTCATCGCCAGGAGTACCCAGATGAGCCAAGATTCGAGACC 550

RESULT   2
AZ399131/c
LOCUS   AZ399131   474 bp   DNA   GSS   03-OCT-2000
DEFINITION 1M0164P14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0164P14 R, DNA sequence.
ACCESSION   AZ399131
VERSION   AZ399131.1 GI:10514203
KEYWORDS   GSS.
SOURCE   house mouse.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 474)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE   Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0164   row: P   column: 14
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 474.
FEATURES   Location/Qualifiers
     source   1..474
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0164P14"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain Xll0-Gold, Tl-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli Xll0-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT   92 a   151 c   131 g   100 t
ORIGIN

Query Match   9.6%;   Score 383.8;   DB 181;   Length 474;
Best Local Similarity 93.7%;   Pred. No. 3.1e-79;
Matches 400;   Conservative   0; Mismatches 27;   Indels   0; Gaps   0;

Qy   1531 ctcaccctcatcgccaagtt 1551
          ||||| |||| |||||||
Db   551 CTCACCCTCATCGCCAAGGGT 571

Plasmid DNA from the normalized library NCI_CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323311, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT   104 a   169 c   178 g   141 t   1 others
ORIGIN

Query Match   7.4%; Score 294.6; DB 95; Length 593;
Best Local Similarity   73.0%; Pred. No. 2.3e-58;
Matches   378; Conservative   0; Mismatches 140; Indels   0; Gaps   0;

```
Qy    841 gggaaaggagctgctgcctgctgtggctgaaggccgttaccagacaatgagtatcctg 900
          ||| || |  ||  | | ||| ||| || || |||||| | |||||||||||| || ||
Db    518 GGCAAGGGACTGACCCATGATCAGCATCAAGGCGCGCTACCAAACCATCACCATCCTG 459

Qy    901 cccatggactatataaggagttgcagaatatgtgaccaaccactaccgcatgctgtgt 960
          || |||||| | | ||| ||| ||||| |||| ||||||| |||||| |||| |||| |
Db    458 CCCATTGAGATGTACAAAGAGTTCGTGACGCACATCACCACTACCTGTGGCTGTGT 399

Qy    961 gccgtgctgagcccgccctcaatgtcaaggcaaggaggagtcgctagtgcactggtt 1020
          || | |||| ||||| ||||  | ||| ||||| | |||| ||| || | ||||||||
Db    398 GCAGCCCTCGAGCCCATCCTCAGTGCCAAGACCAAGAAGACGAGATGGCATCTGCCCTGGTG 339

Qy   1021 cacatcctgcaaagcacaggcaaggccaaggacttcctttcagacatggccatgtcagag 1080
          ||||||||| |  ||||| |||||| | ||||| ||||| || ||||||| ||||||||||
Db    338 CACATCCTGCAGAGCACGCGGCAAGGTGAAGGACTTCCTGACAGACCTGATGATGTCAGAG 279

Qy   1081 gtagacggttcatggagcgggaacacctcatattccgcgaagaaacacgctcgccactaaa 1140
          ||||||  || | | ||| | |  |  || |||| ||| ||||  |||| ||||||||| |
Db    278 GTGGACCGCTGCGGGACAACGACCACCTCATCTTCCGGGAGAACAACACTGCCACCAAG 219

Qy   1141 gccatagaagtatagacgactgattggccagaaatacctcaaggatgccattggggag 1200
          ||| || | || | |   ||  | |||  ||||| |||  ||||||||||||| |||||
Db    218 GCCATTGAGGAGTACCTCAAGCTAGTAGTGGGCCAGAAGTACNTGCAGGACGCCCTAGGTGAG 159

Qy   1201 ttcatccgggctcgtatgaatctgagagagaactgtgaagtagaccccatcaagtcaca 1260
          || ||||||||| ||  | | || |||| || ||| |||| ||||| || | ||| || ||
Db    158 TTCATCAAAGCGCTGTATGGATCAGATGAGAACTGCGAAGTGGATCCCAGCAAGTGCTCG 99

Qy   1261 gcgtccagtcggcagcaccaggccaacctgcggatgtgtgctgtgagttggccctgtgc 1320
          || |||| |||| | |  || || ||||| |||| |||| || ||||| ||||||||||
Db     98 GCCGGTGACCTCCAGACGCACCAGGGCAACCTCAAGATGTGTGCCGAGCTGGCCTTCTGC 39

Qy   1321 aaggtggtcaactccattgcgttcccgagggagct 1358
          ||| ||||||| | |||| | | || |||||||||
Db     38 AAGATCATCAACTCCTACTGTGTCTTCCCACGGGAGTT 1
```

RESULT 4
BE463433/c
LOCUS   BE463433   463 bp   mRNA   EST   27-JUL-2000
DEFINITION   hw23d06.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3183755 3'
             similar to TR:O95174 O95174 NGAP. ; mRNA sequence.
ACCESSION   BE463433
VERSION   BE463433.1   GI:9509206
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 463)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov

```
Qy    423 caaggacaacgccgccggtagataacgtgctgaaacatggatcatagagctcgaga 482
          | |||||||||| |||||| ||| ||| || || |||| ||| |||||| ||||||
Db    427 CCAGGACAACAGCCCGCGAGCTAGATAACGTGCTGAAGCTGATCATAGAGGCTCGAGA 368

Qy    483 gctgccccccaagaagcgatattactgcgagtcctggacgacatgctctatgcacg 542
          |||||||| ||||||| | | ||||| || |||| ||| | ||||| ||| ||||||
Db    367 GCTGCCCCCAAGAAGAGATATTACTGCGAGCTGTGCCCTGGACGACACATGCTGTATGCACG 308

Qy    543 gaccacttcaagccccgctcagctcaggagacactgtctttggggcgagcacttcga 602
          |||||| ||||| |||||| | || |||||| || | || ||| ||| ||||||||| |
Db    307 AACCACTTCAAGCCCCGCTCGCTCAGGAGACACCGTCTTTTGGGGCGAGCACTTTGA 248

Qy    603 gttaacaacctgcctgctcgtccgggcgctgcgcgtgcatctgtaccgtgactcggacaa 662
          |||||| ||| ||| | |||||| || |||  |  | | ||| | |||| |||| |||| |
Db    247 GTTTAACAACCTGCCTGCTCGTCCGGGCGCCCTTCGCGTGCCATCTGTACCGTGACTCAGACAA 188

Qy    663 aaagcggaagaagcaggcagttggcctacgttgcctggctgattcagttgccaccct 722
          ||||| || ||||| |||| | |||| || |||||||||| |||| ||||| ||||| |
Db    187 AAAGCGGAAGAAGCAGGCTGGCCTACGTGCCTGGGCTGTTCCAGTGGCCACCCT 128

Qy    723 ggctgggcgccacttcacagagcagtggtgtacccctgacccgtccaacaggagagtgggg 782
          ||||| ||||||| |||||||||  ||||| | ||| ||||||| ||| |||||||| ||||
Db    127 AGCTGGGCGCCACTTCACAGAGCAGTGTACCCCGTGACCTTGCCGACAGGCAGTGGGG 68

Qy    783 ctctggggtatggctcgggggagagggggggtcaggggcggctcaggggcaaag 842
          |||||| |||| |||| ||| ||||| |||| | | |||||||||||| ||||||
Db     67 CTCTGGGGCATGGCTCGGGGGAGGAGGGGTGGCTCAGGGGCAAGG 8

Qy    843 gaaagga 849
          ||||||
Db      7 GAAAGGA 1
```

RESULT 3
AW779747/c
LOCUS   AW779747   593 bp   mRNA   EST   12-MAY-2000
DEFINITION   hn85f10.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3034699 3'
             similar to TR:O95174 O95174 NGAP. ; mRNA sequence.
ACCESSION   AW779747
VERSION   AW779747.1   GI:7794350
KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 593)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Tel: (301) 496-1550
             Email: Robert_Strausberg@nih.gov
             Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
             Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL, send email to:
             info@image.llnl.gov
             Seq primer: -40UP from Gibco
             High quality sequence stop: 451.
FEATURES   Location/Qualifiers
     source   1..593
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3034699"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
```